NBER WORKING PAPER SERIES

PHYSICIAN GROUP INFLUENCES ON TREATMENT INTENSITY AND HEALTH: EVIDENCE FROM PHYSICIAN SWITCHERS

Joseph J. Doyle Jr.
Becky Staiger

Working Paper 29613
http://www.nber.org/papers/w29613

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2021, Revised January 2025

We thank the editor, Heather Royer, along with Leila Agha, Laurence Baker, David Chan, David Cutler, Phillip Decicca, William Evans, John Friedman, Daniel Hungerman, Anupam Jena, Kirabo Jackson, Ethan Lieber, David Molitor, Mohan Ramanujan, Jonathan Skinner, Tara Templin and seminar participants at ASHEcon and Notre Dame for helpful comments and suggestions. We gratefully acknowledge support from the National Institutes of Health R01 AG41794. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Physician Group Influences on Treatment Intensity and Health: Evidence from Physician Switchers
Joseph J. Doyle Jr. and Becky Staiger
NBER Working Paper No. 29613
December 2021, Revised January 2025
JEL No. I10

ABSTRACT

Treatment intensity varies remarkably across physicians, yet key drivers of this variation are not well understood. Meanwhile, physicians are increasingly working in groups that may influence how they practice. This paper tests whether physicians' group affiliation matters for practice styles and patient health. Using Medicare inpatient claims data, we compare these outcomes before and after physicians switch between groups of varying treatment intensity while remaining in the same hospital to control for practice setting. Event studies show that internists who join groups with higher Medicare payments per patient immediately increase their own treatment intensity, with an elasticity of approximately 0.27; the opposite is found for internists who switch to groups that are less intensive. This change in Medicare spending largely stems from higher-priced services, although an increase in the quantity of care is also affected. We do not detect a change in health outcomes, suggesting that treatment intensity induced by group affiliation may not be productive.

Joseph J. Doyle Jr.
MIT Sloan School of Management
100 Main Street, E62-516
Cambridge, MA 02142
and NBER
jjdoyle@mit.edu

Becky Staiger
2121 Berkeley Way
#5424
Health Policy & Management
School of Public Health
Berkeley, CA 94720
bstaiger@berkeley.edu

A data appendix is available at http://www.nber.org/data-appendix/w29613

1 Introduction

With nearly one in every five dollars spent in the US going to healthcare, there is considerable interest in understanding the underlying drivers of this spending (CMS, 2023). It is widely believed that treatment decisions by physicians are key contributors to healthcare spending, summarized by the saying that the most expensive equipment in healthcare is the pen of the physician. It is also well known that physicians vary remarkably in their treatment intensity across regions, and even within the same hospital (Fisher et al., 2003a; Tsugawa et al., 2017). The sources of this variation are not well understood, however (Berndt et al., 2015; Chan, 2021; Cutler et al., 2019; Epstein and Nicholson, 2009). Recent research has started to fill in this gap, identifying potential sources of variation that include demand-side factors, such as patient illness severity and preferences; and supply-side factors, such as physician practice styles and institutional constraints and incentives (Molitor, 2018; Finkelstein et al., 2016; Cutler et al., 2019).

Meanwhile, the organization of healthcare is undergoing a long-term transformation as physicians increasingly work in physician groups (Capps et al., 2017; Kane, 2017). By 2016, we find that 84% of physicians work in groups instead of as solo practitioners, including 93% of inpatient physicians, and that these shares have been growing over time.[1] This secular trend away from solo practice is explained by a variety of factors that include growing administrative burdens and incentives to coordinate care. Despite considerable discussion of this supply-side shift in healthcare organization, the implications for treatment intensity and patient health merit more attention (Heeringa et al., 2020; Zwiep et al., 2021; Muhlestein and Smith, 2016; Welch et al., 2013). Moreover, a better understanding of group influence on physician practice style would not only inform the implications of the trend towards group practice, but may also shed light on a fundamental question in healthcare productivity: what are the key drivers of practice-style variation across physicians?

This paper tests whether a physician's group affiliation influences their treatment intensity and the health outcomes of their patients. Simply comparing physicians who belong to groups that vary in their practice styles could be misleading due to various forms of endogeneity bias. For example, patients treated by different groups may vary in their illness severity or preferences for treatment, and physicians may influence one another simultaneously, leading to a reflection problem (Manski, 1993). Our empirical strategy aims to circumvent these endogeneity concerns by comparing physicians who switch groups within the same setting and continue to treat the same types of patients. To that end,

[1]This percent is likely an underestimate, given the limitations of generating these estimates based on a 20% sample of Medicare beneficiary medical claims.

we restrict the analysis to physicians who switch groups within the same hospital. We also restrict the analysis to physicians with a specialty of internal medicine (i.e., internists) whose patient mix and departments are similar before and after changes in group affiliation.

We recognize that physicians choose their destination group and may take into account the group's intensity when making that decision. An abrupt and stable change in treatment intensity upon switching groups would suggest that either the group exerts some influence on treatment intensity directly, or a physician may need to switch groups in order to practice in a different way. Either way, such a change points to treatment variation stemming in part from group affiliation.

To make these comparisons, we use Medicare inpatient claims by physicians, detailing the treatment of beneficiaries from 2008 to 2016. We define a group as those who bill using the same (de-identified) tax identification number (TIN) (as in Austin and Baker (2015); Baker et al. (2014b); Ketcham et al. (2007); Welch et al. (2013)). TINs allow us to observe a common payroll function across physicians, which will introduce measurement error to the extent that multiple physician practices bill under the same TIN or bill under multiple TINs. For summary measures of group treatment intensity as measured by Medicare spending, the group billing ID is close to what we are seeking to characterize. Meanwhile, we characterize physician treatment intensity, as well as the intensity of other physicians in their group, using average Medicare spending on each of their patients. We find that a physician who joins a more-intensive group immediately increases her own intensity at the time of the switch. In particular, a one-standard deviation increase in the change in group intensity before and after the switch (approximately 68 log points), results in an 18 log-point increase in physician intensity, translating to an elasticity of 0.27. Switching to more-intense groups raises intensity, while switching to less-intense groups lowers it, with point estimates suggesting that the change induced by joining a less-intense group is somewhat larger.

When we use an alternative approach that simply relates the change in physician treatment intensity to the change in group intensity, the relationship is visible even in this raw comparison. The slope of this relationship implies that 14% of the within-hospital variation in observed intensity across physician groups can be attributed to group-specific factors, while the remaining 86% is attributable to physician-side components.

To explore how patient welfare may be impacted by such group effects, we also evaluate how this change affects several quality-of-care measures, including hospital readmissions and mortality. Despite a change in treatment intensity that scales with the change in the group-intensity measure, we do not detect a change in these health outcomes: we find that readmissions are modestly lower, and mortality is modestly higher, when switching to more-intensive groups, but the estimates are not statistically significant and the visual evidence does not support a sustained change in these outcomes. While other aspects of care can vary with group intensity, group-induced increases in intensity do not appear to be productive.

The identifying assumption underlying the causal interpretation of our results is that physicians who switch groups do not experience contemporaneous shocks correlated with the size of the change in group intensity. For example, patient characteristics could differ across origin and destination groups, or physicians may change their preferred practice style precisely at the time when they switch practices in a way that systematically varies with the change in group intensity. We find that the number of patients treated per quarter, along with observable patient characteristics, are similar before and after the moves, which is consistent with within-hospital changes in group affiliation not affecting the types of patients that physicians treat. We also find no change in the share of patients treated in intensive care, which provides evidence that these internists are not moving groups to work in a critical-care unit where patient characteristics would also change. In addition, we observe that the trajectory of treatment intensity prior to the switch is unrelated to the change in group intensity, which provides additional confidence that the identifying assumption is plausible. Note that we focus on switching physicians and rely on the gradient of the change in group intensity to identify the effects on physician intensity. As a result, the identifying assumptions allow switching physicians to differ from those who do not switch, and we describe differences across these physicians to begin to learn whether the findings for switchers are likely to apply to non-switchers.

We explore further for heterogeneity of effects along three dimensions. First, we search for heterogeneity across moves that are part of a consolidation and those that are not. Non-consolidation moves are closer to the spirit of comparing physicians as the group changes holding constant other characteristics, and we find that the main results are driven by non-consolidation moves. Next, we compare results across different types of physicians. We find that the effects are particularly large for internists, though still present among a non-internist sample. We find similar results at different points in physicians' careers and across different levels of pre-switch intensity. Third, we consider different types of groups. Our main results are robust across groups that vary in size, specialty mix, and origin-group treatment intensity levels. Given the interest in the role of group size, it is noteworthy that switching to a larger group is associated with a modest reduction in healthcare spending.

Last, we explore potential sources for the group's influence on treatment intensity. When we examine whether the elasticity stems from quantity or price, we find that a larger contributor to the spending-intensity relationship stems from providing higher-priced services. In particular, the most common types of claims for internists are "evaluation and management" claims (Cabral et al., 2021); we find that more-intensive groups have such claims that are coded to represent a greater amount of time and effort spent treating the patients, and switching to a more-intensive group leads to a sudden and sustained increase in such coding. Opposite effects are found when physicians switch to less-intensive switches as well. These patterns could reflect greater (less) intensity, a change in coding, or both. Regardless, the results imply that efforts to address spending growth may benefit from reforms aimed at the group level.

The rest of the paper proceeds as follows. Section 2 briefly describes the trend in practicing in groups over time and related literature. Section 3 introduces our empirical framework. Section 4 describes our data and details the sample construction. Section 5 presents the results and Section 6 discusses potential mechanisms. Section 7 concludes.

2 Background and Related Literature

2.1 Variation in Treatment Intensity

Variation in treatment intensity across physicians is remarkable, even among physicians working in similar practice environments and treating similar patients. Wennberg and Gittelsohn (1973) famously showed that tonsillectomy rates varied widely across Vermont towns, launching a large literature documenting remarkable small-area variation in treatment intensity. Potential drivers of this variation include the preferences and training of physicians (Cutler et al., 2019; Epstein and Nicholson, 2009), along with institutional features such as financial incentives, constraints, and practice norms (Clemens and Gottlieb, 2014; Molitor, 2018). Tsugawa et al. (2017) demonstrate that among general internists treating Medicare patients within the same hospitals, physicians at the 90th percentile of spending had 50% higher hospitalization costs compared to the 10th percentile, even after adjusting for patient characteristics. They show that this is relatively larger than the substantial between-hospital, cross-region variation in treatment intensity (Baker et al., 2014a; Barnato et al., 2007; Cutler et al., 2019; Finkelstein et al., 2016; Fisher et al., 2003a,b).

Cutler et al. (2019) explore the black box of "supply-side" drivers of regional variation by using physicians' answers to vignettes of patient cases to identify factors that influence physician behavior. They find that approximately 60% of the variation in end-of-life spending across markets can be explained by whether a physician is classified as a “cowboy” (more aggressive) or a “comforter” (less aggressive), and that physician beliefs regarding the efficacy of therapeutic interventions (not necessarily based on clinical effectiveness) are the key drivers of these differences in intensity, explaining as much as 35% of end-of-life expenditures. While the authors find that group structure (namely, size and single- or multi-specialty practice) explains only a small amount of the variation in physician behaviors, our analysis extends this analysis of group effects by explicitly examining the intensity of the group, which is not captured in the surveys and may be more relevant to a physician's own intensity.

One source of influence on practice styles may be the physician's peers. Epstein and Nicholson (2009) study how residency training and a physician's local peers (in the same hospital or in the same market) might affect a physician's propensity to opt for a C-section during delivery of newborns. The authors find only a very small effect of both training and local C-section rates on a physician's own C-section rate, where residency programs explain approximately 2% of the variation. They conclude that much of the practice variation between physicians is likely due to a physician's beliefs regarding the efficacy or appropriateness of specific treatments. They also document a significant amount of within-region variation in C-section rates, observing that within-market variation is approximately twice as large as variation between markets, although the implications for patient welfare are unclear.[2] Saghafian et al. (2019) and Chan (2016) study emergency-room physicians who practice side by side, finding that physicians who work with faster or higher quality peers tend to perform worse, while holding physicians jointly responsible for their care can reduce a “foot-dragging” form of moral hazard when patients are assigned independently across physicians. Similarly, Silver (2020) finds that emergency-room physicians practicing with more-intense peer groups increase the intensity with which they treat their patients (i.e. spend more time per case), leading to quality of care improvements.

Group-practice affiliation can also affect physician financial incentives via different compensation models, and there is evidence that physicians respond to such financial incentives (McGuire and Pauly, 1991). For example, Clemens and Gottlieb (2014) exploit a regional consolidation of Medicare fees that resulted in significant changes in payment rates across areas to explore the role of financial incentives in physician treatment decisions. They find that a 2% increase in payment rates led to a 3% increase in the provision of care, and that the use of elective procedures was more responsive to this change than non-elective procedures. Alexander (2015) and Alexander and Schnell (2019) use fee changes in Medicaid to

[2]If there is an “optimal” level of C-section frequency, then within-market variation in rates implies that some patients will receive more/less C-sections than recommended, thereby reducing patient welfare. If, however, variation reflects differences in patient preferences or suitability for C-section, then variation could be welfare enhancing.

show that physicians respond to increased payments by increasing their use of C-sections and increasing the number of Medicaid patients they treat, respectively. Cabral et al. (2021) find similar elasticity estimates when examining a payment reform that increased the generosity of Medicaid payments for "evaluation and management" visits.

2.2 Analysis of "Movers"

Closely-related work studies regional variation in treatment and outcomes. Molitor (2018) studies cardiologists who move to a more (or less) intensive area and tests whether this results in a change in treatment intensity. He finds that the environment in which a cardiologist practices (which describes factors such as hospital capacity and productivity spillovers) accounts for 60-80% of the observed variation across hospital referral regions (HRRs) in catheterization rates. Notably, he observes that the effect of a change in intensity in more localized environments (i.e. hospitals) on a physician's own catheterization rate is larger than the effect of a change in intensity at the broader geographic region, suggesting that physician behavior may be especially sensitive to small-area environments. Similarly, Finkelstein et al. (2016) study patient movers among Medicare beneficiaries to decompose regional variation in utilization into demand-side and place-specific, supply-side factors. They find that patient-specific components (such as health and preferences) account for approximately 40-50% of observed variation in healthcare utilization across HRRs. In follow-on work, Badinski et al. (2023) examine physician movers as well as patient movers and find that physicians drive most of the supply-side factors contributing to geographic variation.

This paper offers three main contributions beyond these prior "movers analyses." First, this paper is complementary but in many ways orthogonal to the prior work. While movers analyses allow an examination of the role that market-level influences play in treatment intensity as the environment shifts around the physician, this paper considers how group affiliation affects treatment intensity holding constant the environment where the physician practices. We aim to control for variation-contributing factors that might otherwise change when individuals move across regions, such as area-level resources, patient health, and returns to treatment intensity given the capital and complementary labor mix that changes at the same time as a regional move. As a result, any influence of group affiliation should be all the more remarkable if other aspects of the practice environment are unchanged. Second, by analyzing within-hospital variation, we can distinguish between individual switches compared to switches due to consolidation. Third, we estimate a decomposition of the variation into group-specific and physician-specific factors; the goal of the exercise differs from prior work, as we focus on the role that group affiliation plays as a driver of treatment variation across physicians.

2.3 Physician Group Formation

There are growing incentives for physicians to practice in groups. Market and environmental factors—such as increasing financial burdens associated with medical debt, administrative requirements including quality reporting and documenting meaningful use of health information technology, and policies that generate new incentive structures for more coordinated care—all have prompted a significant shift toward group practices (Harris, 2010; Kane, 2017; Muhlestein and Smith, 2016; Welch et al., 2013). Indeed, most physicians now work in groups. In Medicare claims data that we describe in more detail below, we see a steady decline in the share of physicians practicing in solo practice over time.[3] Figure A.1 (panel a) shows that by 2016, 7% of physicians observed treating patients in the inpatient setting were working in a solo practice, down from 15% in 2008.

Several mechanisms have been proposed to explain how groups can improve performance, including that groups can (1) benefit from economies of scale, such as the incorporation of health IT, and (2) alter compensation models to reward quality of care in addition to the quantity of care provided. Nevertheless, empirical tests of group influence on treatment decisions and patient health have been lagging (for a review, see Zwiep et al. (2021)). Epstein et al. (2010) show that compared to solo practitioners, group practices in obstetrics are better able to match patients to specialists, improving their health. Similarly, for inpatient cardiac care, Ketcham et al. (2007) find that patients treated by physicians in solo practices are less likely to receive invasive procedures and have higher mortality.

Much of the existing research of physician groups focuses on group size. Spending and quality measures have been compared across different-sized groups controlling for practice and physician characteristics using a selection-on-observables approach. For example, Casalino et al. (2014) combine survey data on group-practice size with Medicare quality measures and find that small practices have 30% lower preventable admissions compared to practices with 20 or more physicians. McWilliams et al. (2013) provide more nuanced evidence of larger hospital-based practices providing greater treatment intensity with higher readmission rates, while larger independent physician groups have lower spending levels and higher quality scores. Such comparisons could reflect differences in patient characteristics as group types

[3]Our billing data come from a 20% random sample of Medicare beneficiaries. The share in solo practice is likely even lower, as we may be less likely to observe them, although the trend toward fewer solo practices is evident in the 20% sample.

and sizes vary. In general, the ongoing wave of physician-group formation and consolidation is striking. Even within the hospital setting, Figure A.1 (panel b) shows that average group size in our sample in the first quarter of 2008 was 67 and grew nearly three-fold to 184 by the end of 2016.

Our study adds to the discussion by considering another element of the group environment: group practice intensity. We focus our exploration on the inpatient setting in order to consider short-term, welfare-relevant outcomes including spending, readmissions, and mortality. Exploring group switches within a hospital also allows us to control for fixed attributes of the hospitals where physicians practice.

3 Empirical Framework

3.1 Estimating Group Effects on Physician Behavior

Our goal is to test whether a physician's group matters for how they practice using various measures of treatment intensity. Appendix B includes a simple model of intensity choice that is the result of physician and group effects, taking into account patient characteristics as well. Physicians influence treatment intensity due to their preferences, skill, private (opportunity) costs of administering the care, and their beliefs about the effectiveness of the care (e.g. Ellis and McGuire (1986); Alexander (2015); Clemens and Gottlieb (2014); Cutler et al. (2019)). Groups can influence treatment decisions through productivity incentive structures, billing technology, and the group's relative weighting on profits versus benefits to patients (Dafny, 2005; Song et al., 2020). The end result is a straightforward model of physician intensity in the spirit of Abowd et al. (1999) (hereafter, referred to as "AKM") and Finkelstein et al. (2016) that includes physician (worker) and group (firm) fixed effects.

Abstracting from time-varying characteristics of the environment and patients, the following simplified model of a physician's observed level of intensity in terms of these effects for physician p and group g can be written as:

$$y_{pg} = \alpha_p + \delta_g + \varepsilon_{pg} \tag{1}$$

Where α_p are physician fixed effects; δ_g are group fixed effects; and ε_{pg} are unobserved characteristics that drive variation in intensity, such as patient characteristics. For those physicians who switch to group g' then:

$$y_{pg'} = \alpha_p + \delta_{g'} + \varepsilon_{pg'}$$

Our empirical approach compares physicians before and after they switch groups to physicians that vary in their intensity. The main idea is that physician effects are constant across the switch to a new group. If moves are exogenous, allowing us to ignore the ε terms in expectation, then the change in observed intensity will identify the difference of group effects:

$$E(y_{pg'} - y_{pg}) = \delta_{g'} - \delta_g$$

To estimate these group effects, we relate the change in treatment intensity of switching physicians when they join more- or less-intensive groups. To characterize this change in group intensity, we calculate the degree to which the intensity of the destination group differs from the origin group as:

$$\Delta_{[pmt/pt](p)} = \overline{y}_{d(p),q<0} - \overline{y}_{o(p),q<0} \tag{2}$$

where $\overline{y}_{o(p),q<0}$ and $\overline{y}_{d(p),q<0}$ are the average Medicare payment-per-patient (pmt/pt) of the other physicians in the origin and destination groups, respectively, calculated in the four quarters, q, prior to the switch, where the quarter of the switch is normalized to zero. In terms of notation, $\Delta_{[pmt/pt](p)}$ is defined specifically for each switching physician; however, going forward we omit the p in the subscript for simplicity. As described more fully below, results are nearly identical when we employ estimates of group intensity using an empirical Bayes shrinkage estimator to account for measurement error in $\Delta_{pmt/pt}$ that may arise from small samples.

Note that from Equation 1 the change in group environment represents a change in average physician effects (from other physicians) and group effects as we average over physicians to construct $\Delta_{pmt/pt}$:

$$\Delta_{pmt/pt} = \overline{\alpha}_{d(-p),q<0} - \overline{\alpha}_{o(-p),q<0} + \delta_{d(p),q<0} - \delta_{o(p),q<0} \tag{3}$$

To define the share of the variation in group intensity that stems from physician effects vs. group effects, simply divide both sides by $\Delta_{pmt/pt}$:

$$\begin{aligned} Share_g =& (\delta_{d(p),q<0} - \delta_{o(p),q<0})/\Delta_{pmt/pt} \\ Share_p =& (\overline{\alpha}_{d(-p),q<0} - \overline{\alpha}_{o(-p),q<0})/\Delta_{pmt/pt} \end{aligned} \tag{4}$$

As in Finkelstein et al. (2016), we use an event-study approach in which physician p switches from origin group o to destination group d to recover the average effect of group intensity on a physician's own intensity. We discuss any contamination of the estimates that might arise from using staggered treatment timing below. In this empirical strategy, the jump in a physician's intensity at the time of the switch identifies the extent of the influence of the group environment on a physician's own intensity. Using the definitions in Equation 4 and the timing of the switch, the AKM model can be re-written for event time r as:

$$\begin{aligned} y_{pg} =& \alpha_p + \delta_{o(p),q<0} + \mathbb{1}(r>0)(\delta_{d(p)} - \delta_{o(p)}) + \varepsilon_{pg} \\ =& \alpha_p + \delta_{o(p),q<0} + \mathbb{1}(r>0) Share_g * \Delta_{pmt/pt} + \varepsilon_{pg} \end{aligned} \tag{5}$$

where $\mathbb{1}(r > 0)$ is an indicator for the post-switch period.

When we bring this model to the data, we can include controls for time-varying characteristics of the environment and patients, and our estimating sample includes non-switchers to help estimate the relationships associated with these controls. Our estimating equation models physician p's treatment intensity (or quality of care), y, in group g and hospital h during calendar-quarter t as:

$$y_{pght} = \alpha_p + \beta_{ht} + \sum_{q=-10}^{10} \gamma_q \mathbb{1}\{Q_{pt} = q\} + \sum_{q=-10}^{10} \theta_q \mathbb{1}\{Q_{pt} = q\} \times \widetilde{\Delta}_{pmt/pt} + \lambda X_{pt} + \varepsilon_{pght} \tag{6}$$

where α_p is a physician-episode fixed effect (and "episode" refers to the period of at least nine quarters before and after a switch to a new group at $q = 0$), which controls for tenure in the data when we move to longer time horizons and the panel becomes unbalanced.[4] β_{ht} represents hospital-specific calendar year-quarter fixed effects to control for hospital-specific trends that occur contemporaneously with or around the physician's switch. In some analyses, we include X_{pt}, which represents average patient characteristics measured at the physician-quarter level; and ε_{pght} is an error term that we assume to be mean zero and mean independent of the event-time indicators, their interaction with relative group intensity, and included patient characteristics.

The remaining elements of the empirical model trace the outcomes of interest in the quarters to and from a switch. We are specifically interested in how the difference in intensity across the destination and origin groups, $\Delta_{pmt/pt}$, affects the physician's behavior around the switch. Thus, the main coefficients of interest are the θ_qs.

As is typical in interaction models, we de-mean $\Delta_{pmt/pt}$, represented as $\widetilde{\Delta}_{pmt/pt}$ to ease interpretation

[4] Molitor (2018) uses HRR-level fixed effects instead of physician-level fixed effects in order to test for selection among moving physicians, namely that they are systematically different from baseline migrants in the same HRR. We will explore how those who switch groups differ from those who do not switch below.

of the estimates. In practice, $\Delta_{pmt/pt}$ has a mean that is close to zero. We set all indicator variables for the quarter relative to the switch to 0 for the non-switching cohort.

The pattern we find in the event studies suggests that a more parsimonious model that improves precision is also informative. That is, we estimate a pre-post version of the event study in which we define an indicator, Post Switch, as equal to one for all quarters $q \in [1, 10]$ and zero otherwise, in addition to an indicator, $\mathbb{1}\{\text{Qtr=0}\}$, which is equal to one when $q = 0$ and 0 otherwise to allow for a transition quarter when the physician is practicing in both groups. This consolidated model takes the form:

$$\begin{aligned} y_{pght} =& \alpha_p + \beta_{ht} + \delta_1 \mathbb{1}\{\text{Post Switch}\} \times \widetilde{\Delta}_{pmt/pt} + \delta_2 \mathbb{1}\{\text{Qtr=0}\} \times \widetilde{\Delta}_{pmt/pt} \\ & + \delta_3 \mathbb{1}\{\text{Post Switch}\} + \delta_4 \mathbb{1}\{\text{Qtr=0}\} + \lambda X_{pt} + \varepsilon_{pght} \end{aligned} \tag{7}$$

The object of interest is δ_1, which tests the effect of the change in group intensity on physician intensity in the post-switch period relative to the pre-period, $q \in [-10, -1]$

3.2 Inference

We compute two-way clustered standard errors at the physician and group levels to incorporate correlation within and between these two attributes. Because $\Delta_{pmt/pt}$ is a generated regressor, we also report confidence intervals when using a bootstrap procedure that incorporates the variability due to the calculation of $\Delta_{pmt/pt}$.

3.3 Identification and Interpretation

The goal of our exercise is to compare physicians who change groups, holding constant the setting and the types of patients they treat. As a result, ideal switches are those when a physician switches groups but does not change her role or department, continues to treat the same types of patients, and there is little disruption to the destination group so that we can characterize its practice style in the pre-switch period.

The identifying assumption that allows us to attribute changes in physician intensity to the influence of group affiliation is that $\Delta_{pmt/pt}$ is exogenous. That is, in the absence of group effects, trends in physician treatment intensity would have evolved in ways that are unrelated to the change in group intensity experienced by the physicians. A feature of the event-study approach is that we can observe whether this is the case prior to the move. An absence of pre-trends in treatment intensity related to $\Delta_{pmt/pt}$ provides some evidence in support of the identifying assumption. A violation of this assumption would include physicians experiencing a shock to their preferences at the time of a switch that is not due to group influences.

In our setting, the post-switch slope created by the θ_qs in $q > 0$ is also informative. As Molitor (2018) notes, an immediate jump in θ_q followed by a relatively flat slope in the estimates of θ_qs for $q > 0$ is consistent with group norms and policies driving a sudden, one-time adaptation of the physician's intensity towards that of the group environment. Conversely, an increasing slope (following an immediate jump at $q = 0$) may reflect a more long-term, adaptive group effect. As described in the Appendix, a jump at the time of the move identifies the share of the change in group intensity that stems from group effects as opposed to physician effects.

Our setting features staggered episodes, which allows us to control for calendar-time effects. Sun and Abraham (2020) note that in such settings, estimates may not be a straightforward average across individuals if there is anticipation or heterogeneous treatment effects. We also show that the event study estimates using the pooled, staggered events provide a good summary of the set of event studies that are estimated separately for episodes defined by switches that happen in the same calendar quarter.

Groups with different intensities may also differ along other observable and unobservable dimensions, such as patient characteristics, that could challenge our causal interpretation of the θ_qs. In an attempt to control for such changes, we restrict our analysis to physicians who remain in the same hospital before and after switching groups. However, even after restricting our analysis to within-hospital group switches, physicians may treat different types of patients following a group switch (Chang and Obermeyer, 2020). To investigate this concern, we test for balance of patient characteristics before and after the physicians switch groups. Specifically, we estimate Equation 6, replacing measures of intensity and quality of care on the left-hand side with several key patient characteristics that have been linked to differing levels of treatment intensity and health outcomes. Related, we investigate whether departmental changes affect the interpretation of the results. More-intensive groups may differ along other dimensions as well, such as physician training and beliefs. We view our estimates as a test of whether group affiliation matters for treatment intensity and health outcomes.

A related concern is that a physician may choose to move to a group with higher (or lower) intensity in order to change their treatment intensity, such as physicians starting to taper their practice in preparation for retirement. Here, we again evaluate any presence of pre-trends to help us investigate whether behavior changes in anticipation of the switch. We also note that a sudden change in treatment style at the time of the switch suggests that physicians are constrained in their behavior until the move occurs, which would imply that group affiliation matters for treatment intensity even if physicians choose a group because it is a better match for their preferred intensity level. For retirement influences in particular, we directly estimate effects of switches for physicians of different ages.

In contrast to other studies of movers that focus on regional variation in intensity, our measure of origin-group intensity may reflect the physician's own behavior. Specifically, the switching physician may influence the practice intensity of her peers in the origin group. Given that our main explanatory variable is the difference in treatment intensity across the destination and origin groups, we run several checks to ensure that this potential source of endogeneity is not driving our results. First, we estimate models where we use the destination-group intensity as the main explanatory variable of interest rather than the change in group intensity. This forsakes the useful variation in the shock to group intensity that comes from variation in origin-group levels, but it relies on a potentially more exogenous measure of the shock to the practice environment. Second, we report a set of results that flexibly control for origin-group intensity levels.

4 Data and Sample Description

4.1 Data

Our primary data are traditional (fee-for-service) Medicare claims from 2008 to 2016. To measure physician treatment intensity, we rely on claims for a 20% random sample of beneficiaries in the Carrier file where payments to physicians are recorded. In traditional Medicare, payments for physician services are made on a fee-for-service basis; physicians can increase their reimbursement for a given patient by increasing the services they provide or by selecting more expensive services. The claim includes lines-of-service coded using the Healthcare Common Procedure Coding System (HCPCS), analogous to commercial Current Procedural Terminology (CPT) codes. With these codes, we investigate whether the types of claims change after a switch and begin to consider changes in coding behavior.

Importantly, these data also include a billing identifier (ID), the (de-identified) Tax Identification Number, and we identify groups based on physicians billing under the same ID (Austin and Baker, 2015; Baker et al., 2014b; Ketcham et al., 2007; Welch et al., 2013). There are at least two potential limitations when relying on such a billing ID to characterize the environmental intensity of a physician's group. First, the ID may represent a much larger organization, and the other physicians in the group may not exert as much influence as those working in the same team as a smaller unit, such as within a clinic (Welch et al., 2013); alternatively, a particularly large group may bill under more than one ID (Capps et al., 2018). Second, given that the Carrier file represents a 20% random sample of beneficiaries, we are likely not capturing all physicians associated with a given group, which introduces measurement error as well. To the extent that this measurement error is larger for small groups, we will explore robustness of the results across groups of varying sizes.

We measure each group's intensity as the average per-physician reimbursement-per-patient for all physicians in the group except the switching physician, across the four quarters prior to the switch quarter, weighted by the number of patients that a physician treats.[5] In this way, physicians' contributions to the intensity of the group environment are representative of how active they are in the group. Our results are robust to non-weighted measures of average intensity. These data also include patient characteristics, including age, race, and sex, and because the data are longitudinal we are able to observe claims for the beneficiary before and after an admission.

To carry out the empirical strategy, we limit claims to services performed in the hospital in an inpatient setting, excluding treatment that is given in other settings (such as the emergency department) in order to avoid introducing confounding from switches across departments. This restriction also allows us to merge the physicians' claims records to the 100% Inpatient files to identify the hospital associated with a given stay and the hospitals where physicians work. We also use the Inpatient files to record the admission and discharge dates associated with that hospital stay in order to calculate length of stay, 30-day readmission rates, and the number of major procedures associated with a given stay. These data also provide additional information on diagnoses.[6]

We use the National Plan and Provider Enumeration System (NPPES) data set to obtain additional information about physicians, including gender and specialty, and to differentiate between physicians and other medical professionals. We also use the CMS Physician Compare database to obtain information on physician experience (in years), based on the year they graduated from medical school. We match

[5]We average the quarterly intensity measures by the number of patients treated each quarter. We use the claim's summary payment amount measure, available in the Carrier files, which is the sum of payments made by CMS to the physician and the beneficiary. Beneficiary payments tend to be negligible on average ($< 0.1\%$ of the total payment), and thus we take these payments to characterize the amount a physician receives from CMS.

[6]We merge Carrier and Inpatient claim records based on de-identified patient ID and dates of service. According to conversations with the Research Data Assistance Center, an advantage of this approach over relying solely on the place-of-service codes in the Carrier files is that it more accurately captures hospital stays.

85% of our final treated sample of physicians to this database. In robustness checks where we search for treatment effect heterogeneity by years of experience, we focus our analysis on this sub-sample. Last, we use American Hospital Association survey data to identify general acute care hospitals and observe the share of patients at the hospital covered by Medicare, which we use in a robustness check.

4.2 Outcomes

Our outcomes are estimated at the physician-quarter level and are intended to capture measures of treatment intensity and quality of care. The main treatment intensity measure is again reimbursement per patient (the measure used to characterize the groups). We then test whether a switching physician provides more services per patient, measured by the number of line items filed each with its own HCPCS code, and whether they provide higher-priced services, measured by payments per HCPCS.

We estimate effects on additional measures of intensity at the patient rather than physician level: measures of the number of major procedures and the length of stay. Because these measures are linked to the entire hospitalization, and thus not necessarily attributable to the switching physician, we include them as a representation of broader treatment intensity (attributing procedures to *all* physicians who had corresponding Carrier claims associated with that hospitalization).

Next, we include several measures intended to capture changes in the quality of care provided. First, we calculate a physician's 30-day readmission rate as the share of all patients the physician treated in a given quarter who had a readmission within 30 days of the discharge date. Second, we calculate a physician's 30- and 365-day mortality rates as the share of hospitalizations in which the patient died within 30 or 365 days of admission. The mortality measures stem from vital statistics records, so we observe mortality regardless of whether it occurs in a hospital or not.

These measures are commonly used to evaluate the quality of care provided. Thirty-day readmission is used by CMS as a quality measure.[7] The 30-day mortality rate in particular is included in Hospital Compare data as a measure of hospital quality (Doyle et al., 2019). Note that we are attributing these readmission and mortality rates to physicians who are *not necessarily* listed as the attending physician on the hospitalization record, but instead have a corresponding carrier line item during the hospitalization. This approach allows us to estimate a readmission and mortality measure for each physician in our sample, though it potentially deviates from more conventional approaches of attributing

[7] We do not differentiate between unplanned 30-day readmissions, which are penalized by CMS, and planned readmissions in order to measure total resources used.

readmissions/mortality to the attending physician on record.

4.3 Sample Construction

Our study sample is comprised of two physician cohorts: physicians who switch groups ("switchers"), and physicians who never switch groups ("non-switchers"), whose primary function is to increase the precision with which we can estimate and control for hospital- and calendar-level secular trends. In addition to focusing on inpatient treatment by internists, we make a number of additional sample restrictions to implement our estimation strategy, as shown in Table A.1. Because we examine the effect of group environment on physician intensity, we attribute physicians to exactly one group per quarter, where group assignment is determined by the billing identifier associated with at least 90% of their claim line items (represented by HCPCS) for which they file for reimbursement in that particular quarter. On average, physicians associate 92% (SD: 22%) of their HCPCS with a particular billing ID in any given quarter.[8] Of the 553,721 physicians in our starting sample, we were able to attribute 552,420 to one group per quarter.

We define a switching episode for each physician by identifying a period of at least nine consecutive quarters during which the physician belongs to a given origin group for at least four consecutive quarters, switches to a destination group in a "switch quarter," and belongs to that destination group for at least four consecutive quarters thereafter. By this definition, switching physicians can have multiple episodes. We observe 72,426 physicians who ever switch, and 83,870 switching episodes; each switching physician is associated with an average of 1.16 (SD: 0.40) episodes. Figure A.2 plots the share of HCPCS associated with a given origin or destination group for physicians in our final sample, in the quarters relative to the switch. As is evident from the figure, there is a transition quarter at the time of the switch ($q = 0$), during which physicians transition out of their origin group to the destination group.

Non-switcher physicians include any physician who is observed to be attributed to only one group throughout the study period, which we similarly refer to as their "episode" for the sake of consistency. By this definition, we flag 321,963 never-switching physicians, included for an average of 15.9 (SD: 13.3) quarters during our study period. The other physicians who were dropped at this step were in multiple groups but did not meet our nine consecutive quarter restriction (to be included in our switcher cohort).

To focus on within-hospital variation in treatment intensity, we further restrict the sample to physi-

[8] In approximately 7% of treated physician-quarters outside the switch quarter, physicians with an internal medicine specialty are attributed to groups associated with less than 90% of HCPCS in that quarter because they had more than 90% of HCPCS associated with a single group in the surrounding quarters.

cians practicing within one general acute care hospital during each episode. Note that switcher physicians may switch hospitals at some point during the study period, as long as it does not occur contemporaneously with a group-switch episode. We attribute each physician to exactly one hospital per quarter by assigning them to the hospital associated with the plurality of their HCPCS in a given quarter.[9] On average, physicians associate 62% (SD: 40%) of their HCPCS with a particular hospital in any given quarter.

In order to be included in the final sample, both the origin and destination groups must exist in the four quarters prior to the physician's switch, as this is the relevant time period for measuring the level of intensity. Additionally, in order to calculate the change in environmental intensity, which is estimated based on the average intensity of the *other* physicians in the group, at least one other physician (in addition to the switching physician) must belong to the origin group. This restriction limits the analysis in two ways: First, we cannot observe origin groups in which the switching physician was the solo practitioner. Instead, we evaluate how effects vary by size of the origin and destination groups to see whether there is a relationship between origin-group size and our main results. Second, destination groups that do not exist in the pre-switch period are excluded from the sample, which excludes any group that forms in the post-switch period. This restriction focuses our analysis on changes in a physician's own intensity level due to a change in group intensity that arises from already-established group environments. Finally, to calculate group intensity (and to mitigate measurement error), we require that each origin and destination group treat at least 10 patients per quarter. We show that our results are robust to different cutoffs and to the use of empirical Bayes estimates to characterize the intensity level of a group.

After imposing these restrictions, we have 162,433 non-switching physicians, 30,887 of whom have a specialty of internal medicine. As detailed in Table A.1, we observe 13,883 switching physicians (14,487 physician-episodes), including 3,108 physicians with a specialty of internal medicine (3,242 physician-episodes).

Because we specify that switching physicians belong to an origin group for at least four quarters before the switch, and a destination group for at least four quarters after the switch, we have an unbalanced panel when we examine outcomes beyond those quarters. Physician fixed effects (detailed in our model in Section 3) control for any systematic, time-invariant differences between physicians that

[9]In the instance of a tie (i.e. a quarter in which the physician has equal HCPCS across multiple hospitals), we default to the general acute care hospital, and remaining ties are broken at random; these ties occur for approximately 4% of physician-quarters.

are in a given group for exactly four quarters and those that are in a given group for more than four quarters.

4.4 Descriptive Statistics

Figure 1 plots the distribution of group intensity (panel a) and physician intensity (panel b), as well as the relationship between the two (panel c). Notably, the standard deviations of log payment-per-patient are quite large; a one-standard-deviation increase in overall group intensity is 0.56, and 0.47 for within-hospital variation in group intensity. This large degree of variation in group intensity is interesting in its own right, and it is also useful empirically for our identification strategy. As with group intensity, the standard deviations of physician intensity are quite large; a one standard deviation increase in physician intensity is 0.61 overall and only slightly smaller at 0.51 when measured within groups. When we consider all physicians including non-internists, these standard deviations increase by approximately 30%, which we report in Figure A.3. This large degree of variability is both remarkable and in line with prior literature (Epstein and Nicholson, 2009; Tsugawa et al., 2017).

Physician intensity is positively and strongly correlated with the intensity of their peer colleagues. Without any additional adjustments, panel (c) of Figure 1 shows that physicians who belong to an origin group that has a 100 log point higher group intensity (approximately 2 standard deviations) have a 42 log point higher intensity level themselves.[10] This correlation does not account for any endogeneity that might be associated with both the group intensity and the physician's intensity, such as a physician's preference for practicing in a group similar to their preferred level of intensity, or features of more-intensive groups (such as increased physical or human capital) that influence a physician's intensity. We aim to control for these factors in our analysis below.

In the quarter prior to a switch within the analysis sample, physicians in the analysis sample work in hospitals with an average of 14 groups (SD: 12) that have at least one internist member. Table 1 provides context for the types of groups and physicians that are considered using the empirical strategy. Column (1) describes the full sample, including the physicians who do not switch. Next, column (2) describes the group and physician characteristics for the switching physicians; and columns (3) and (4) report the characteristics of physicians who switch to more- or less-intensive groups, respectively. The

[10]For computational reasons, for non-switchers we calculate the intensity of the group using all physicians (a leave-in mean); thus, own intensity is highly correlated with group intensity, particularly in smaller groups. Because we don't believe this to be informative, but rather reflective of a mechanical relationship that we avoid in our leave-one-out means, we exclude non-switchers from Figure 1c.

first row shows that the main explanatory variable of interest—the change in treatment intensity—is -53 log points for those switching to less intensive groups and 37 log points for those switching to more-intensive groups. This again demonstrates the striking heterogeneity in group intensity even within the same hospital. Figure A.4 plots the distribution of the relative change in group intensity as a histogram, showing a standard deviation of 0.68.

In levels, the average origin-group intensity measured by average reimbursement-per-patient is $275; for switchers, this origin-group intensity is somewhat lower at $218.[11] Physicians who leave origin groups for destination groups that are less intensive come from groups that have a relatively high pre-switch intensity ($252), while physicians switching to more-intensive groups leave groups that are slightly lower intensity with a mean of $191. Not surprisingly, the opposite trends are found for destination groups. Whenever we cut the sample by the direction of the change in group intensity, the origin group is on average relatively more intensive when the physician is moving to a lower-intensity group and vice versa, as expected due to the nature of the subsample selection.

Switchers tend to join larger groups, as indicated by the number of patients (303 vs. 254) and number of physicians (130 vs 93). The movement towards larger groups, and how it corresponds to changes in group intensity, is something we explore in more detail below. While some switches may be due to a reorganization, one-quarter are solo moves where only one physician is switching that quarter, and the average (median) number of physicians who switch out of a given group at one time is 4 (2);[12] for the most part, these are not acquisitions of origin groups.

In terms of specialty mix, the share of physicians who are internists is 0.50 in the origin groups and 0.34 in the destination groups, partially reflecting the secular trend of physicians moving from smaller to larger, more multi-specialty groups. Another way to consider specialty mix is by the number of diagnostic categories observed, defined by the hierarchical ICD-CM-9 and ICD-CM-10 sections, such as "Diseases of the Circulatory System." Groups in our data average 9 categories prior to a switch, increasing to 11 following a switch regardless of the direction of treatment intensity.[13]

Panel C reports physician characteristics. Average reimbursement among switchers in the pre-period is somewhat lower compared to all physicians, as their origin groups are also relatively less intensive. Among switchers, physicians tend to move to groups that are more similar to their pre-switch intensity: those moving to less-intensive groups tend to be approximately 11% less intensive than their origin-

[11] While we measure group intensity for switchers as a leave-one-out mean, we calculate group intensity for non-switchers as simply the overall average, inclusive of the index physician, for computational reasons.

[12] Note that these are switching physicians who meet our definition of switchers.

[13] Groups tend to serve patients from a large geographic area, averaging 25 (s.d. = 28) patient ZIP codes per year-quarter.

group peers prior to the switch ($28 less compared to a mean of $252) and 37% more intensive than their destination group prior to the switch. Similarly, those moving to more-intensive groups tend to be 6% more intensive than their origin-group peers and 25% less intensive than their destination-group peers prior to the switch. Given that the level of intensity varies across physicians who choose different destination groups, we rely on the identification assumptions described above and reflected in the event studies below to estimate the causal effect of the change in group intensity on own intensity. Nevertheless, these differences motivate our exploration below of results by different levels of origin-group treatment intensity, as well as different levels of physician treatment intensity prior to the switch.

Switchers are somewhat less likely to be male than all physicians (62% versus 69%). Switcher physicians have slightly fewer years of experience at 21 years compared to an overall average of 23. As a whole, physician characteristics are fairly similar across different types of switches.

5 Main Results

5.1 Balance Checks

Our goal is to examine the effects of an exogenous shock to a physician's group affiliation while maintaining the same setting and types of patients. All group switches in our analysis sample are made by internists who switch between groups within the same hospital.

A first check on whether these moves preserve the practice environment is to consider balance on observable patient characteristics. Table 2 reports estimates of the difference-in-differences model represented in Equation 7 using patient characteristics on the left-hand side as the outcomes of interest. We observe no meaningful changes in the majority of the characteristics. One exception is patient age, although the estimated effects do not exhibit a clear jump at the time of the switch (Figure A.5, panel b). Moreover, the magnitude is modest relative to the mean. We find no relationship between changes in group intensity and race and sex, nor in the composition of admitting diagnoses. Event study figures for patient characteristics and the top 10 admitting diagnosis shares support the lack of systematic changes in these characteristics (Figures A.5-A.7).

As a summary measure of predetermined patient illness severity, we calculate both a predicted mortality and predicted inpatient spending measure using patient demographics and diagnoses in Medicare claims data in the year prior to the hospitalization. To calculate predicted mortality, we first used a linear model to estimate the relationship between an indicator for whether a patient died in 2012 or 2013 and patient age (in vigintiles), sex, race, and indicators for diagnoses recorded in 2012, with 2012 being the midpoint of our study period among patients not included in our analysis sample. This measure of illness severity is similar across different moves of increasing or decreasing intensity. Despite a strong relationship with actual mortality, we find no meaningful difference in predicted mortality, which we interpret as evidence against physicians treating demonstrably different patients following a group switch (Figure A.8). We construct predicted inpatient spending in a similar way, also finding no relationship between switching groups and predicted spending (Figure A.9). Finally, we find no change in the number of patients treated across the switch. Due to evidence that patient composition doesn't change along these dimensions, our main results do not control for patient characteristics, although we report results with these controls in the appendix as discussed below.

5.2 Group Affiliation and Physician Reimbursement

Figure 2 presents the main event-study results. The horizontal axis represents the quarters relative to the group switch. The points represent the θ_qs estimated using Equation 6: the difference in log payment-per-patient in the quarters leading up to and lagging away from a switch, scaled by the difference in treatment intensity between the destination and origin groups. Panel (a) shows that the relationship between physicians' treatment intensity and the eventual change in group intensity is relatively flat and approximately 0 prior to the switch. We see a small jump in treatment intensity in the quarter of the switch, followed by a substantial increase that remains steady for the following 10 quarters. Specifically, we observe that the relationship is relatively steady at approximately 0 log points prior to the switch quarter, and then rises to an elasticity estimate of just under 30 log points after the switch, staying relatively constant in the post-switch quarters. Confidence intervals are similar when we bootstrap the standard errors (Figure A.10) to take into account that the measure of the change in group intensity is a generated regressor.

Panels (b) and (c) report the same event studies for physicians who join more (less) intensive groups. The results are somewhat noisier, as expected given the smaller sample sizes and nature of the restricted variation, but the direction of the change in intensity is symmetric across the two types of moves. Physicians joining more-intensive groups see an elasticity with respect to the change in group intensity of approximately 20 log points. For those who join less-intensive groups, the change in own intensity elasticity appears sustained at approximately 40 points lower than pre-switch intensity.[14]

[14]The lack of symmetry in the point estimates points to the possibility of a model that departs from additive separability

Recent discussions surrounding contamination in the estimates of event-study models that include staggered events have prompted a re-examination of traditional estimation methods (Sun and Abraham, 2020; De Chaisemartin and d'Haultfoeuille, 2022). In particular, there is a concern that the event-study coefficients plotted need not represent an average of effects across the staggered events. To determine whether this concern is relevant in our setting, we estimate our model separately for each event period defined by the calendar quarter of the switch and compare our results to the average of these many event studies. We identify 28 switching cohorts with an average of 116 (SD: 51) physicians per cohort. In Figure A.11, we plot the estimated θ_qs from each cohort-specific episode (a "scatter plot of event studies"), and find that the average of these θ_qs is similar to our main, pooled estimates. These results indicate that our estimated effects closely approximate the average of effects across the cohorts.

Table 3 reports estimates from the more parsimonious pre-post model given by Equation 7. We estimate that an increase in group intensity by 1 (or, 100 log points, which is similar to a 2 standard deviation increase in $\Delta_{pmt/pt}$) results in an approximately 27 log-point increase in a physician's own intensity. This is substantial, although significantly smaller than the raw correlation described above that implied an elasticity of closer to 0.42.

5.3 Decomposition of Group and Physician Effects

As noted in Section 3.1 and in Finkelstein et al. (2016), the AKM model allows us to decompose group-intensity variation into components that are attributed to the physician (such as preferences and beliefs about treatment) and groups (such as group management). Equation 5 shows that the share of group variation attributed to the group effects is the slope of the relationship between $\Delta_{pmt/pt}$ and the jump in physician's treatment intensity at the time of a switch. This can be flexibly estimated using a bin scatter plot of the average change in treatment intensity among switchers across bins of $\Delta_{pmt/pt}$.

Figure 3 carries out this exercise using vigintiles of the treatment variable (the difference in treatment intensity across the destination and origin groups). Note that there are no controls in this specification; we simply bin the data according to vigintiles of the change in group intensity and plot the average change in physicians' intensity. The relationship with the change in group intensity is fairly linear, which is reassuring, as (1) it is consistent with the additively-separable AKM model, and (2) it suggests that the results are not driven by only a small portion of the distribution which might be related to shocks

of the AKM model. We investigate this further below when we plot the change in physician intensity versus the change in group intensity, where linearity suggests that additive separability may be reasonable in this setting.

in the unobservables related to the switch. The coefficient of 0.14 can be interpreted to mean that approximately 14% of the observed variation in treatment intensity across groups can be attributed to a group-specific component, leaving 86% of the variation to be attributed to a physician-specific component among internists.

The jump in intensity at the time of the switch from origin to destination group in the event study also provides an estimate of $S_{group}(g, g')$. We interpret the somewhat smaller magnitude of the slope in Figure 3 compared to the size of the jump in Figure 2 (approximately 0.27), as evidence that the hospital-quarter controls in our main results account for general trends in intensity over this time period. In any event, these different approaches demonstrate that group affiliation is a significant contributor to treatment-intensity variation across physicians.

5.4 Price vs. Quantity

While average Medicare reimbursement-per-patient is a useful and policy-relevant summary measure of a physician's practice style, we can decompose the sources of our main results into changes in the quantity of services provided and the price of those services. Namely, do physicians who switch to higher intensity groups provide higher-priced services, more services, or both?

Figure 4 reports event studies where the outcomes are measures of the lines of claims per patient (HCPCS/pt), and payment-per-HCPCS as a measure of price-per-service. Here, we see a stronger effect with respect to the price of services provided compared to the quantity of services provided. This is confirmed in Table 3, which reports an elasticity of 0.048 for our quantity measure, while the payment-per-HCPCS increases with an implied elasticity closer to 0.19.

Taken together, the estimates suggest that physicians' group affiliation matters: when their group intensity increases, physicians increase their own intensity of treatment, particularly in terms of performing tasks that have a higher reimbursement and a modest increase in the number of services provided per patient, with no clear evidence of changes in underlying patient population.

5.5 Heterogeneity

5.5.1 Individual switches versus group consolidation

The goal of our empirical exercise is to compare physicians whose group affiliation changes while holding constant their role and practice setting. As noted above, useful switches in the context of our empirical exploration would involve one (or a handful) of physicians who are switching out of a group at a particular point in time. In contrast, examining changes in physician behavior after the consolidation of a group practice may confound changes in group affiliation with any disruption in norms that take place at the same time, as the practice style of the destination group will simultaneously be in flux.

To test the robustness of our main results to restricting our sample to this "ideal" population of switchers, we define a consolidation-based switch as one where all physicians switch out of the group such that the origin group ID no longer exists after that quarter. When a physician switches in this context, they nearly always make up less than 25% of the destination group's physicians, so this definition appears to identify instances when the switching physician's group is acquired. Table 4 shows that physicians who switch groups that are not induced by a consolidation are similar to the main results. Those who are the only physician switching from an origin group–more akin to our ideal thought experiment–exhibit a larger elasticity of 0.37, and those who move in the same quarter with at least one other physician in a non-consolidation move (the majority of switches) exhibit an elasticity of 0.21.

In contrast, for those who appear to be in a group that is acquired by another group, we see virtually no response. The lack of an effect for switches that are (plausibly) a result of a consolidation is consistent with measurement error (the pre-switch intensity of the destination group may no longer reflect the group's practice style, which may change following the acquisition). It is also consistent with the idea that physicians may need to seek out another group in order to practice according to their preferred style.

5.5.2 Heterogeneity Across Specialties

Our main results focus on internists for two reasons. First, in principle, restricting our analyses to one specialty allows us to control for the type of care provided. Second, internal medicine is the most common specialty among the switching physicians in our sample. To determine if group effects are internist-specific or whether they are more universal, we also explore whether groups affect treatment intensity and health outcomes among other types of physicians.

When we estimate our model on the pooled sample of all physicians of all specialties, we find similar patterns of the effects of group intensity on physician intensity, though the magnitudes of the effects are smaller (Table A.2).[15] When we implement the decomposition exercise and plot the average change in physician intensity for all physicians against the associated change in group intensity vigintiles the slope

[15] We observe similar effects of group intensity among cardiologists (another common specialty), though the effects are smaller and less precise.

is 0.08. We interpret this set of results as evidence supporting the conclusion that a change in group environmental intensity applies to a range of physicians, with larger effects found for internists.

5.5.3 Relationship with Physician Age

The causal interpretation of our results depends on the assumption that there is not a contemporaneous shock at the time of a switch, such as a change in patient characteristics or preferred practice style, that might also explain a change in treatment intensity. For example, physicians switching groups at particular points in their career could reflect changing priorities that may also generate the patterns we observe among our physician switchers. More specifically, consider older physicians who seek to scale back their workload and who may switch to lower-intensity groups at the same time as they change their preferred practice style. Figure A.12 plots the main coefficients of interest (θ_qs) by quartiles of physicians' years of experience. All four quartiles show a relatively flat pre-trend followed by a sustained increase in intensity, including for those with the most experience. We view this evidence as supportive of our main identifying assumption.

5.6 Robustness Checks

The main results are robust to a wide range of alternative characterizations of treatment intensity. Our primary measure of the change in group intensity is the difference between average intensity among physicians in the destination group and average intensity of other physicians in the origin group, both measured in the four quarters before the switch. However, if peers influence one another as we hypothesize, then the origin-group intensity will reflect the influence of the switching physician. As an alternative measure of group intensity that circumvents this potential bias, we estimate our model using the destination-group intensity as the treatment of interest rather than the change in intensity. Using this alternative definition, we estimate effects that are very similar to our primary results (Figure A.13). We also expect that a physician may have a larger influence on her peers in small groups, yet results are similar across different group sizes, a topic we return to below. Further, when we estimate group intensity using only non-switchers, we find similar results (Table A.3).

Table 1 shows that switchers to less-intensive groups are less intensive than their colleagues, and that switchers to more-intensive groups are more intensive than their colleagues. This motivates our exploration of effects across different levels of origin-group intensity. Here again we find similar results regardless of the quartile of origin group intensity (Figure A.14) or of the quartile of physicians' pre-switch intensity (Figure A.15).

Our measure of group treatment intensity is likely to be more precisely estimated for larger groups, leaving open the possibility that we may be estimating intensity with some degree of error among smaller groups. To mitigate the possibility of bias arising from small sample sizes, we measure group intensity across the year prior to the switch, restricting our analyses to groups that treat at least 10 patients per quarter. After imposing this restriction, the average number of patients per group is over 200, which should provide a relatively precise measure of the mean intensity. We test the robustness of our 10-patient-per-quarter minimum inclusion criteria by estimating Equation 6 using 5-patient and 20-patient cutoffs, finding effect sizes that are very similar to our main estimates (Figure A.16). We interpret this similarity as evidence that our results are largely robust to any measurement error that may enter our measures of group intensity. We also verify that the results are nearly identical when we implement Bayesian shrinkage to characterize the treatment intensity of the groups (Figure A.17, Table A.3).

If physicians switch departments at the same time as they switch groups, this could affect the types of patients they treat. To keep the empirical setting as constant as possible across switches, we restrict our analysis to treatment in the inpatient setting, so that switches to a group whose members treat patients predominantly in the emergency department (for example) are not included. We also consider environmental changes within the broader inpatient setting, such as a switch to a group whose member physicians treat patients primarily in a critical care unit. Using type-of-service codes that describe the location of the service, we find that the average physician in our sample has a relatively small share of claims (3%) in a critical care setting.Indeed, if we estimate our model on the subset of physicians who never treat patients in a critical care unit, the results are largely unchanged from our main results (Figure A.18).[16] While the use of intensive care units is an endogenous outcome of interest, these results are reassuring that changes in treatment setting are not driving the main results. Moreover, Figure A.19 demonstrates the robustness of our results to including controls for patient characteristics.

Our measure of group intensity is based on claims data generated by Medicare patients only, which may not accurately characterize a physician's practice style across all (Medicare and non-Medicare) patients. When we use data from the American Hospital Association to estimate the model separately for hospitals that vary in their share of Medicare patients, the results are qualitatively similar across these categories (Table A.4). In fact, we find a larger responsiveness in the hospitals with the lowest Medicare-patient share, which demonstrates heterogeneity across hospitals that is contrary to this type

[16]We find similar results when we exclude physicians whose change in the share of claims located in the intensive care unit (ICU) is greater than or equal to the top and bottom 5% of the distribution of the change in the share of ICU claims.

of measurement error contaminating our main estimates.

We further explore the robustness of the spirit of our findings to alternative measures of physician intensity (holding constant our definition of group intensity as average payment per beneficiary across all member physicians). Table A.5 reports that payment measures made to the physician aggregated to the quarterly level (not the patient level) yield similar results.[17]

5.7 Effects on Health Outcomes

To evaluate welfare implications of the impact of group intensity on physician intensity, we next consider health outcomes. If changes in physician intensity following a switch to a more or less intensive group result in better patient outcomes, then this may suggest that group influence is productive. If, instead, group changes result in worse or no change to patient outcomes, then group influence may not reflect clinical improvements to a physician's practice. Figure 5 reports the results when estimating our main specification on measures of readmissions and mortality. Despite a sudden and sustained change in how much a physician bills Medicare and the number (and intensity) of the procedures they perform, the figures show a relatively flat relationship between the timing of the switch and patient outcomes. The mortality coefficients are both positive and negative in the post period, and within a fairly narrow range.[18]

To gain precision, we again use a pre-post model and report the results in Panel B of Table 3. This approach yields estimated effects indicating that a one-unit increase in post-switch group intensity is associated with a 0.18 percentage point reduction in 30-day readmission, or about 0.7% of the mean of 24%. For 30-day mortality, an increase in group intensity from low spending to high spending (a change of 1) is associated with a 0.2 percentage point increase in mortality compared to a mean of approximately 10% (i.e. approximately 2% of the mean). For one-year mortality, the coefficient represents a 0.4 percentage point increase compared to a mean of 30%. The estimates are reasonably precise: the lower bound of the confidence interval for one-year mortality is -0.3%, or 1% of the mean.[19] With small positive point estimates and a lack of visual evidence of a sustained change in these quality measures, it appears that an increase in treatment intensity induced by joining a more-intensive group

[17]When we consider more aggregated measures of patient treatments due to decisions made by all physicians who treat the patient, including major procedures and length of stay, we do not find a relationship with group intensity. Rather, the change in practice is found when we focus on the care directly provided by the switching physician.

[18]We also do not detect a relationship with mortality when we consider all specialties.

[19]Using an alternative measure of group intensity—quarterly payment per physician, rather than quarterly payment per patient per physician—we find similar results in magnitude and the bounds of the confidence interval do not include any reduction in mortality.

is not associated with improved outcomes for patients.[20]

Perhaps the lack of an effect on major outcomes is not surprising given the magnitudes of the spending differences, which are fairly small on their own. That said, the estimated increases in the price per claim and the number of line-items billed are not trivial. Another reason we may not see that the change in treatment intensity translates to a change in major health outcomes is that the physician is only one of many who might treat any given patient in the hospital, so her own effect may be diluted. Rather, we view the results as consistent with group affiliation affecting treatment intensity that generates a change in Medicare spending with no detectable effect on health outcomes.

We also conduct a more targeted examination of quality measures by exploring the effect of a switch on mortality among patients aged 85 and older who have higher mortality rates overall. Figure A.20 plots the θ_qs obtained from estimating the model on the share of patients older than 85 who die within 30 days and one year. As in our main mortality measures, we observe no meaningful change in mortality surrounding the group switch, reinforcing the primary findings of a largely null effect of increased intensity on welfare.

6 Interpretation and Mechanisms

6.1 Coding Intensity

One mechanism that could drive an increase in billing is through changes in coding behavior. For example, Dafny (2005) documented hospitals "upcoding" patients to higher-paying diagnosis-related groups (DRGs) following a policy that changed reimbursement for certain DRGs. The most common types of HCPCS among internists are evaluation and management (E&M) visits; they account for the vast majority of the claim line-items we observe. A feature of these types of HCPCS is that they have different billing levels depending on patient complexity, and this complexity requires documentation. Some groups may be more efficient in coding visits to increase revenue.

E&M visits in an inpatient setting have three levels of increasing intensity. Table A.6 reports CPT codes and associated average reimbursement associated with each code. In our data, a physician who conducts a level one inpatient initial E&M visit (of approximately 30 minutes duration; CPT code 99221) is reimbursed $97.40 on average. In comparison, a physician who conducts a level three inpatient

[20]When we restrict the sample to attending physicians who direct the care and are more likely to have an effect on health outcomes, we continue to find a positive but statistically insignificant relationship with mortality. In particular, for 30-day mortality we find a coefficient of 0.011 (s.e.=0.006).

initial E&M visit (of approximately 70 minutes duration; CPT code 99223) is reimbursed $194.89 on average.[21]

Figure 6 plots the post-switch change in log volume of these E&M visits, by level, scaled by the change in group intensity between the destination and origin group as in our main specification (Equation 6). We observe no meaningful change in level 1 visits. We observe a significant increase in level 2 and 3 E&M visits, with a slightly larger increase in level 3 visits. This implies that in addition to performing more HCPCS, physicians are billing at a higher intensity as they switch to more-intensive groups.

Additionally, we consider two other measures of coding changes: the number of distinct diagnoses recorded per patient, and the use of diagnoses that signal higher patient complexity.[22] In both exercises, we do not find compelling evidence that switching to more or less intensive groups has a meaningful impact on coding practices. These findings provide additional reassurance that the characteristics of the patients a switching physician treats do not change substantially across group switches.

Taken all together, the above exercises suggest that physicians who switch to more (less) intense groups spend more (less) time with patients, or they change their coding habits to reflect this change in time spent. Either way, the higher-priced line items in the physician's claim do not appear to be the result of a change in patient complexity, and we do not find an improvement in patient health outcomes. This suggests that this group-induced change in intensity may not be productive.

6.2 Changes in other group characteristics: Size and Specialty Mix

When physicians switch groups, group intensity is not the only component of their practice environment that changes. As noted above, we observe that physicians on average switch to larger groups with increasing variance in their specialty mix. Such group attributes may affect treatment intensity as well, as group size and treatment intensity are negatively correlated in our analysis sample (Figure A.21). When we account for both the change in intensity and the change in size in our main estimating model, our primary estimate for intensity remains essentially unchanged (Table A.7 and Figure A.22), suggesting that the relationship between group size and intensity does not explain the impact of group intensity on physician intensity. Interestingly, we find that physicians who switch to larger groups appear to have

[21] See https://emuniversity.com/Page2.html and https://emuniversity.com/Page4.html for more details.

[22] As a proxy for diagnoses that signal higher complexity, we used a mapping of diagnosis codes to diagnostic-related groups that are characterized as conditions with complications or comorbidities; or with major complications or comorbidities.

modestly lower treatment intensity afterward.[23]

A related characteristic of moving to a larger group is the potential for a change in the mix of specialties. If a physician moves to a group with more-intensive specialists, the physician may be more likely to confer with (and refer patients to) these new peers. Figure A.23 shows that our main results hold for physician switchers regardless of the change in internist share, with some evidence of larger effects for those who switch to groups where the internist share increases, consistent with stronger peer effects rather than referral effects.[24]

7 Conclusion

As physicians increasingly work in group practices to reduce their own financial burden, legal exposure, and resource requirements, a natural question that arises is how group affiliation affects a physician's own practice style and, ultimately, patient health outcomes. Meanwhile, small-area variation in treatment intensity has received considerable attention as a potential context in which to identify strategies to reduce waste. A better understanding of the sources and consequences of small-area variation may inform more effective payment reforms aimed at increasing the efficiency of healthcare spending.

We find that when physicians switch groups within the same hospital, their treatment intensity moves in the direction of the group they join. Among internists, the elasticity of own intensity with respect to group intensity is approximately 0.26. While most of the cross-group variation in intensity is due to physician factors, group factors affect treatment intensity as well, especially for internists. After observing changes in physician treatment intensity that scale with the change in group intensity, we find no corresponding, sustained change in patient health outcomes as measured by readmissions and mortality.

The results have a number of limitations. First, we estimate the influence of group intensity on switchers, who may be more (or less) influenced by group affiliation compared to those who remain in the same group. Second, physician preferences could change at the same time as a switch, such as instances when a change in physician circumstances leads them to make a move. If this is the case, then the sudden and permanent change evident in the event studies suggests that physicians have to wait

[23] To explore how a change in group intensity might be influenced by a change in group size further, we found that the results were similar when we inspected them separately by quartile of the change in group size, across quartiles of origin group size, and quartiles of destination group size.

[24] For more information on this variation, Figure A.24 shows the distribution of the share of physicians who are internists in switcher physicians' origin and destination groups, along with the distribution of the change in this share.

until they make the move before they can realize their new level of preferred treatment intensity. This pattern is consistent with the group exerting influence in how treatment intensity is determined.

Third, given our use of a leave-out estimator of group intensity, we cannot estimate our model on physicians who switch from a solo practice to a larger group. However, the vast majority of physicians practice in groups, and this share is increasing over time. In addition, our results are similar when we consider different sizes of groups that physicians leave and join.

Fourth, our results speak specifically to group influence in an inpatient setting, where there may be less physician discretion for treatment. This approach controls for time-invariant characteristics of the practice setting, but the results are less likely to apply to the outpatient setting, a subject for future research in this area.

Despite these limitations, it appears that group affiliation has a sizeable effect on physician treatment intensity. This helps inform the sources of the remarkable amount of variation across physicians, even those practicing similar roles in the same practice setting. As a result, efforts to restrain healthcare spending may benefit from changing incentives and constraints at the group level.

References

John M. Abowd, Francis Kramarz, and David N. Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999. doi: https://doi.org/10.1111/1468-0262.00020. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0262.00020.

Diane Alexander. Does physician pay affect procedure choice and patient health? evidence from medicaid c-section use. Working Paper WP-2017-7, FRB of Chicago, July 2015. URL https://ssrn.com/abstract=2984189.

Diane Alexander and Molly Schnell. The impacts of physician payments on patient access, use, and health. Working Paper 26095, National Bureau of Economic Research, July 2019. URL http://www.nber.org/papers/w26095.

Daniel R. Austin and Laurence C. Baker. Less physician practice competition is associated with higher prices paid for common procedures. *Health Affairs*, 34(10):1753–1760, 2015. doi: 10.1377/hlthaff.2015.0412. URL https://doi.org/10.1377/hlthaff.2015.0412.

Ivan Badinski, Amy Finkelstein, Matthew Gentzkow, and Peter Hull. Geographic Variation in Healthcare Utilization. *NBER Working Paper*, (31749), 2023.

Laurence C. Baker, M. Kate Bundorf, and Daniel P. Kessler. Patients' preferences explain a small but significant share of regional variation in medicare spending. *Health Affairs*, 33(6):957–963, 2014a. doi: 10.1377/hlthaff.2013.1184. URL https://doi.org/10.1377/hlthaff.2013.1184. PMID: 24889944.

Laurence C. Baker, M. Kate Bundorf, Anne B. Royalty, and Zachary Levin. Physician practice competition and prices paid by private insurers for office visits. *JAMA*, 312(16):1653–1662, 10 2014b. ISSN 0098-7484. doi: 10.1001/jama.2014.10921. URL https://doi.org/10.1001/jama.2014.10921.

Amber E. Barnato, M. Brooke Herndon, Denise L. Anthony, Patricia M. Gallagher, Jonathan S. Skinner, Julie P. W. Bynum, and Elliot S. Fisher. Are regional variations in end-of-life care intensity explained by patient preferences? a study of the us medicare population. *Med Care*, 45(5):386–393, May 2007. doi: 10.1097/01.mlr.0000255248.79308.41.

Ernst R. Berndt, Robert S. Gibbons, Anton Kolotilin, and Anna Levine Taub. The heterogeneity of concentrated prescribing behavior: Theory and evidence from antipsychotics. *Journal of Health Economics*, 40:26–39, 2015. doi: 10.1016/j.jhealeco.2014.11.003.

Marika Cabral, Colleen Carey, and Sarah Miller. The impact of provider payments on health care utilization: Evidence from medicare and medicaid. *NBER Working Paper*, (29471), 2021.

Cory Capps, David Dranove, and Christopher Ody. Physician practice consolidation driven by small acquisitions, so antitrust agencies have few tools to intervene. *Health Affairs*, 36(9):1556–1563, 2017. doi: 10.1377/hlthaff.2017.0054. URL https://doi.org/10.1377/hlthaff.2017.0054. PMID: 28874481.

Cory Capps, David Dranove, and Christopher Ody. The effect of hospital acquisitions of physician practices on prices and spending. *Journal of Health Economics*, 59:139 – 152, 2018. ISSN 0167-6296. doi: https://doi.org/10.1016/j.jhealeco.2018.04.001. URL http://www.sciencedirect.com/science/article/pii/S016762961730485X.

Lawrence P. Casalino, Michael F. Pesko, Andrew M. Ryan, Jayme L. Mendelsohn, Kennon R. Copeland, Patricia Pamela Ramsay, Xuming Sun, Diane R. Rittenhouse, and Stephen M. Shortell. Small primary care physician practices have low rates of preventable hospital admissions. *Health Affairs*, 33(9):1680–1688, 2014. doi: 10.1377/hlthaff.2014.0434. URL https://doi.org/10.1377/hlthaff.2014.0434. PMID: 25122562.

David C. Chan. Teamwork and moral hazard: Evidence from the emergency department. *Journal of Political Economy*, 124(3):734–770, 2016. doi: 10.1086/685910. URL https://doi.org/10.1086/685910.

David C. Chan. Influence and information in team decisions: Evidence from medical residency. *merican Economic Journal: Economic Policy*, 13(1):106–137, 2021. doi: 10.1257/pol.20180501.

Cindy Y. Chang and Ziad Obermeyer. Association of Clinical Characteristics With Variation in Emergency Physician Preferences for Patients. *JAMA Network Open*, 3(1):e1919607–e1919607, 01 2020. ISSN 2574-3805. doi: 10.1001/jamanetworkopen.2019.19607. URL https://doi.org/10.1001/jamanetworkopen.2019.19607.

Jeffrey Clemens and Joshua D. Gottlieb. Do physicians' financial incentives affect medical treatment and patient health? *The American Economic Review*, 104(4):1320–1349, 2014. ISSN 00028282. URL http://www.jstor.org/stable/42920741.

CMS. National health expenditure fact sheet. Technical report, Centers for Medicare and Medicaid Services, 2023. URL https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nhe-fact-sheet.

David Cutler, Jonathan S. Skinner, Ariel Dora Stern, and David Wennberg. Physician beliefs and patient preferences: A new look at regional variation in health care spending. *American Economic Journal: Economic Policy*, 11(1):192–221, February 2019. doi: 10.1257/pol.20150421. URL https://www.aeaweb.org/articles?id=10.1257/pol.20150421.

Leemore S. Dafny. How do hospitals respond to price changes? *American Economic Review*, 95(5):1525–1547, December 2005. doi: 10.1257/000282805775014236. URL https://www.aeaweb.org/articles?id=10.1257/000282805775014236.

Clément De Chaisemartin and Xavier d'Haultfoeuille. Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. Technical report, National Bureau of Economic Research, 2022.

Joseph Doyle, John Graves, and Jonathan Gruber. Evaluating Measures of Hospital Quality: Evidence from Ambulance Referral Patterns. *The Review of Economics and Statistics*, 101(5):841–852, 12 2019. ISSN 0034-6535. doi: 10.1162/rest_a_00804. URL https://doi.org/10.1162/rest_a_00804.

Randall P. Ellis and Thomas G. McGuire. Provider behavior under prospective reimbursement: Cost sharing and supply. *Journal of Health Economics*, 5:129–151, 1986. doi: https://doi.org/10.1016/0167-6296(86)90002-0.

Andrew J. Epstein and Sean Nicholson. The formation and evolution of physician treatment styles: An application to cesarean sections. *Journal of Health Economics*, 28(6):1126 – 1140, 2009. ISSN 0167-6296. doi: https://doi.org/10.1016/j.jhealeco.2009.08.003. URL http://www.sciencedirect.com/science/article/pii/S0167629609000770.

Andrew J. Epstein, Jonathan D. Ketcham, and Sean Nicholson. Specialization and matching in professional services firms. *The RAND Journal of Economics*, 41(4):811–834, 2010. ISSN 07416261. URL http://www.jstor.org/stable/25746055.

Amy Finkelstein, Matthew Gentzkow, and Heidi Williams. Sources of Geographic Variation in Health Care: Evidence From Patient Migration. *The Quarterly Journal of Economics*, 131(4):1681–1726, 07 2016. ISSN 0033-5533. doi: 10.1093/qje/qjw023. URL https://doi.org/10.1093/qje/qjw023.

Elliott S. Fisher, David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F.L. Lucas, and Etoile L. Pinder. The implications of regional variations in medicare spending. part 1: The content, quality, and accessibility of care. *Annals of Internal Medicine*, 138(4):273–287, 2003a. doi: 10.7326/0003-4819-138-4-200302180-00006. URL https://www.acpjournals.org/doi/abs/10.7326/0003-4819-138-4-200302180-00006. PMID: 12585825.

Elliott S. Fisher, David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F.L. Lucas, and Etoile L. Pinder. The implications of regional variations in medicare spending. part 2: Health outcomes and satisfaction with care. *Annals of Internal Medicine*, 138(4):288–298, 2003b. doi: 10.7326/0003-4819-138-4-200302180-00007. URL https://www.acpjournals.org/doi/abs/10.7326/0003-4819-138-4-200302180-00007. PMID: 12585826.

Gardiner Harris. More doctors giving up private practices. *New York Times*, Mar 2010. URL https://www.nytimes.com/2010/03/26/health/policy/26docs.html.

J Heeringa, A Mutti, MF Furukawa, A Lechner, KA Maurer, and E. Rich. Horizontal and vertical integration of health care providers: A framework for understanding various provider organizational structures. *Journal of Integrated Care*, 1(20):1–10, 2020. doi: 10.5334/ijic.4635.

Carol K. Kane. Updated data on physician practice arrangements: Physician ownership drops below 50 percent. Policy research perspectives, American Medical Association, 2017.

Jonathan D. Ketcham, Laurence C. Baker, and Donna MacIsaac. Physician practice size and variations in treatments and outcomes: Evidence from medicare patients with ami. *Health Affairs*, 26(1):195–205, 2007. doi: 10.1377/hlthaff.26.1.195. URL https://doi.org/10.1377/hlthaff.26.1.195.

Charles F. Manski. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542, 1993. ISSN 00346527, 1467937X. URL http://www.jstor.org/stable/2298123.

Thomas G McGuire and Mark V Pauly. Physician response to fee changes with multiple payers. *Journal of Health Economics*, 10(4):385 – 410, 1991. ISSN 0167-6296. doi: https://doi.org/10.1016/0167-6296(91)90022-F. URL http://www.sciencedirect.com/science/article/pii/016762969190022F.

J. Michael McWilliams, Michael E. Chernew, Alan M. Zaslavsky, Pasha Hamed, and Bruce E. Landon. Delivery System Integration and Health Care Spending and Quality for Medicare Beneficiaries. *JAMA Internal Medicine*, 173(15):1447–1456, 08 2013. ISSN 2168-6106. doi: 10.1001/jamainternmed.2013.6886. URL https://doi.org/10.1001/jamainternmed.2013.6886.

David Molitor. The evolution of physician practice styles: Evidence from cardiologist migration. *American Economic Journal: Economic Policy*, 10(1):326–56, February 2018. doi: 10.1257/pol.20160319. URL http://www.aeaweb.org/articles?id=10.1257/pol.20160319.

David B. Muhlestein and Nathan J. Smith. Physician consolidation: Rapid movement from small to large group practices, 2013–15. *Health Affairs*, 35(9):1638–1642, 2016. doi: 10.1377/hlthaff.2016.0130. URL https://doi.org/10.1377/hlthaff.2016.0130. PMID: 27605644.

Soroush Saghafian, Raha Imanirad, and Stephen J. Traub. Do physicians influence each other's performance? evidence from the emergency department. Working Paper RWP19-018, May 2019.

David Silver. Haste or Waste? Peer Pressure and Productivity in the Emergency Department. *The Review of Economic Studies*, 88(3):1385–1417, 08 2020. ISSN 0034-6527. doi: 10.1093/restud/rdaa054. URL https://doi.org/10.1093/restud/rdaa054.

Lina D. Song, Soroush Saghafian, Joseph P. Newhouse, Mary Beth Landrum, and John Hsu. The impact of vertical integration on physician behavior and healthcare delivery: Evidence from gastroenterology practices. Working Paper RWP20-031, October 2020.

Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects, September 2020. URL http://economics.mit.edu/files/14964.

Yusuke Tsugawa, Ashish K. Jha, Joseph P. Newhouse, Alan M. Zaslavsky, and Anupam B. Jena. Variation in Physician Spending and Association With Patient Outcomes. *JAMA Internal Medicine*, 177(5):675–682, 05 2017. ISSN 2168-6106. doi: 10.1001/jamainternmed.2017.0059. URL https://doi.org/10.1001/jamainternmed.2017.0059.

W. Pete Welch, Alison Evans Cuellar, Sally C. Stearns, and Andrew B. Bindman. Proportion of physicians in large group practices continued to grow in 2009–11. *Health Affairs*, 32(9):1659–1666, 2013. doi: 10.1377/hlthaff.2012.1256. URL https://doi.org/10.1377/hlthaff.2012.1256. PMID: 24019373.

John Wennberg and Alan Gittelsohn. Small area variations in health care delivery. *Science*, 182(4117): 1102–1108, 1973. ISSN 0036-8075. doi: 10.1126/science.182.4117.1102. URL https://science.sciencemag.org/content/182/4117/1102.

Terry Zwiep, San (Hilalion) Ahn, Jamie Brehaut, Fady Balaa, Daniel I. McIsaac, Susan Rich, Tom Wallace, and Husein Moloo. Group practice impacts on patients, physicians and healthcare systems: a scoping review. *BMJ Open*, 11, 2021. doi: https://dx.doi.org/10.1136/bmjopen-2020-041579.

8 Tables and Figures

Table 1: Summary Statistics, Internal Medicine

Measure	(1) All	(2) Switchers	(3) $\Delta_{pmt/pt} < 0$	(4) $\Delta_{pmt/pt} > 0$
A. Change in group intensity				
$\Delta_{pmt/pt}$		-.03	-.53	.37
B. Group characteristics				
Pmt/pt per physician				
Origin	275	218	252	191
Destination		222	163	270
Num. patients				
Origin	239	254	227	276
Destination		303	323	286
Num. physicians				
Origin	163	93	83	102
Destination		130	132	128
Share(Internists)				
Origin	.54	.5	.49	.51
Destination		.34	.34	.34
Num. diagnostic categories				
Origin	9	9	9	9
Destination		11	11	11
C. Physician characteristics				
Pmt/pt	245	210	224	202
(Pre) pmt/pt v. origin	-30	-7	-28	11
(Pre) pmt/pt v. destination		-10	61	-68
Share(male)	.69	.62	.66	.6
Mean years experience	23	21	22	20
Total Physician-Episodes	34129	3242	1459	1783

This table presents summary statistics for physicians and groups in our main empirical sample. Column (1) represents overall unadjusted averages/shares for the entire study sample. Column (2) reports averages and shares for the switching physicians, as defined in the text. Column (3) reports statistics for physicians whose destination group is less intense than their origin group, while column (4) represents physicians whose destination group is more intense than their origin group. Pmt/pt = Payment per patient and represents intensity of practice, as defined in the text. Group characteristics are calculated in the four quarters prior to the switch for switchers, and over all quarters for non-switchers. They represent the average for the group in a hospital and a year-quarter. Internists have a specialty of Internal Medicine. Number of diagnostic categories refers to hierarchical ICD-CM-9 and ICD-CM-10 sections. Years experience is calculated as 2016 minus the year of graduation from medical school.

Table 2: Balance Table, Patient Characteristics

Patient Characteristic	Mean	DiD Estimate	SE	p-value
Mean age	75	.263	.111	.018
Share(male)	.43	0	.004	.949
Share(White)	.82	.003	.003	.253
Share(Black)	.13	-.002	.002	.324
Share(claims) by admitting-diagnosis				
Circulatory	.34	.001	.003	.716
Symptoms, Signs	.14	-.001	.004	.859
Respiratory	.1	0	.002	.897
Genitourinary	.07	0	.002	.881
Digestive	.06	-.002	.002	.258
Endocrine, etc.	.06	.001	.002	.768
Injury, Poisoning	.05	.001	.001	.678
Infectious, Parasitic	.04	-.001	.001	.412
Musculo, Connective	.03	0	.001	.751
Blood	.03	0	.001	.693
Mean predicted mortality	.13	.001	.001	.164
Mean number of patients per quarter	11.75	-.147	.432	.734

This table describes characteristics of the patients treated by physicians in our main empirical sample. Column (1) represents overall unadjusted averages/shares of the patient attribute. Column (2) reports the coefficient on $\Delta_{pmt/pt}$*Post Switch obtained from estimating our main difference-in-differences (DiD) specification with the patient characteristic as the outcome in the regression model. Columns (3) and (4) report the standard error and p-value, respectively, associated with the DiD estimate using standard errors that are two-way clustered at the physician and group levels. Share(claims) in major International Classification of Disease (ICD) categories reports the share of admitting diagnoses that are associated with the top 10 most common hierarchical ICD-CM-9 and ICD-CM-10 sections. Predicted mortality is calculated using demographics and all diagnoses recorded in the year prior to a given hospitalization. Number of patients per quarter refers the number of patients treated by the physician.

Table 3: Difference-in-Differences

Panel A. Treatment Intensity	(1) Ln(Pmt/Pt)		(2) Ln(Pmt/HCPCS)		(3) Ln(HCPCS/Pt)	
$\Delta_{pmt/pt}$*Post Switch	0.266***	(0.037)	0.193***	(0.029)	0.048***	(0.008)
$\Delta_{pmt/pt}$*Qtr=0	0.067**	(0.023)	0.050**	(0.018)	0.012	(0.007)
Constant	5.268***	(0.001)	4.310***	(0.000)	1.318***	(0.000)
Dep. Var. Mean	5.267		4.309		1.318	

Panel B. Quality of Care: Readmissions and Mortality	(4) Share(30-Day Readm)		(5) Share(30-Day Mort)		(6) Share(365-Day Mort)	
$\Delta_{pmt/pt}$*Post Switch	-0.0018	(0.0028)	0.0020	(0.0024)	0.0042	(0.0038)
$\Delta_{pmt/pt}$*Qtr=0	-0.0033	(0.0053)	0.0043	(0.0043)	0.0104	(0.0074)
Constant	0.2420***	(0.0001)	0.1012***	(0.0001)	0.3076***	(0.0001)
Dep. Var. Mean	0.2423		0.1015		0.3078	
Observations	529465		529465		529465	

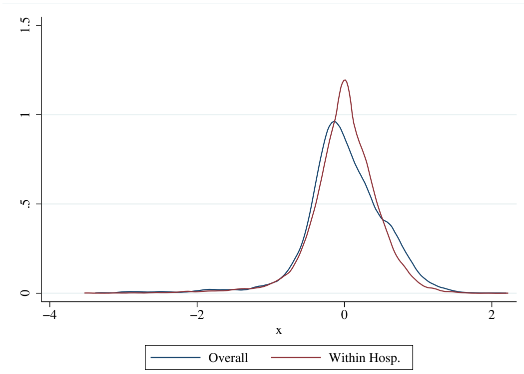
This table reports estimated coefficients from Equation 7 for physicians in our main empirical sample. Post Switch is an indicator variable that is equal to 1 for all quarters $\in [1, 10]$. Fixed effects are included for physician-episode and hospital-year-quarter. Standard errors in parentheses are two-way clustered at the physician and group levels. Omitted category is an indicator for quarters $\in [-10, -1]$. HCPCS is the Healthcare Common Procedural Coding System code recorded as the specific line item in a given claim. "Pmt" abbreviates payment, "Pt" abreviates patient. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Difference-in-Differences, by Switch Type

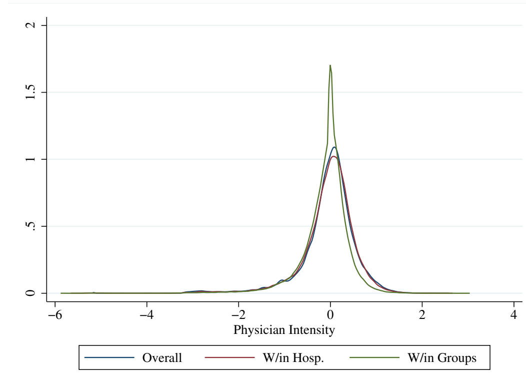
	(1) Ln(Pmt/Pt)	(2) Ln(Pmt/Pt)	(3) Ln(Pmt/Pt)
$\Delta_{pmt/pt}$*Post Switch	0.372***	0.214***	0.022
	(0.047)	(0.053)	(0.035)
$\Delta_{pmt/pt}$*Qtr=0	0.074*	0.116**	-0.035
	(0.030)	(0.044)	(0.054)
Constant	5.271***	5.274***	5.274***
	(0.000)	(0.001)	(0.000)
Observations	486847	503157	482701
Dep. Var. Mean	5.270	5.272	5.273
Switch Type	Solo	> 1 switcher, Non-Consolidation	Consolidation
Num. Physicians	841	1784	595

All models include fixed effects for physician-episode and hospital-year-quarter. Standard errors in parentheses are two-way clustered at the physician and group levels. Consolidation switches are identified as switches where the origin group no longer exists. Pmt/pt = payment per patient. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

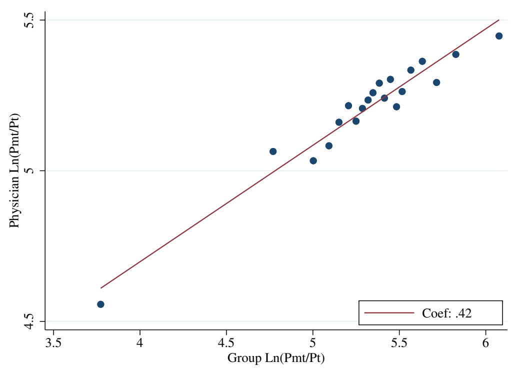
Figure 1: Physician Intensity and Group Intensity



(a) Group Intensity, Internists



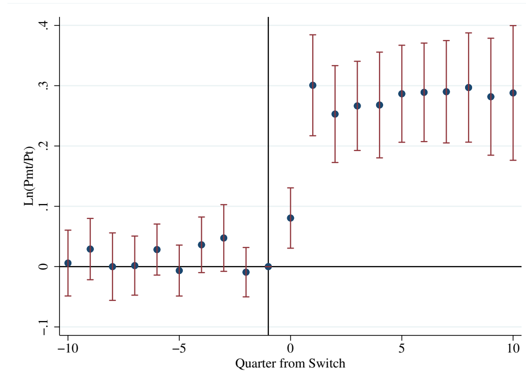
(b) Physician Intensity, Internists



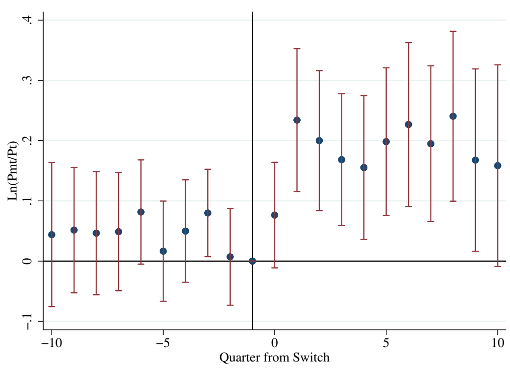
(c) Physician Intensity v. Group Intensity

This figure documents trends in physician and group intensity among the physicians in our main empirical sample, as measured by average log reimbursement per physician per quarter, as well as the relationship between the two. Group intensity is calculated as the average physician intensity calculated across all quarters. Panels (a) and (b) plot the variation in (demeaned) physician and group intensity overall, within hospitals, and within groups (for physicians only), for switchers, non-switchers, and all other out-of-sample physicians associated with in-sample groups. Within-hospital and within-group intensity is demeaned using the hospital- and hospital-group specific averages, respectively. The standard deviation for overall and within-hospital group intensity is 0.56 and 0.47, respectively. The standard deviation for overall, within-hospital, and within-group intensity for physicians is 0.61, 0.59, and 0.51, respectively. Panel (c) plots the relationship between physician intensity and group intensity for switchers. We identify vigintiles of group intensity, and collapse the physician-quarter-level data to averages at these vigintiles, plotted here. The coefficient and standard error are obtained from regression of un-collapsed (i.e. physician-quarter level) physician intensity on group intensity, with no additional controls.

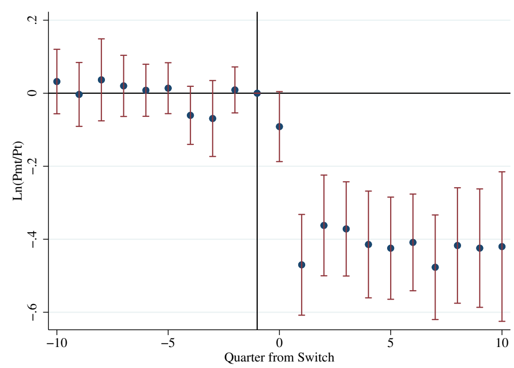
Figure 2: Physician Treatment Intensity Relative to a Change in Group Intensity



(a) Pooled $\Delta_{pmt/pt}$



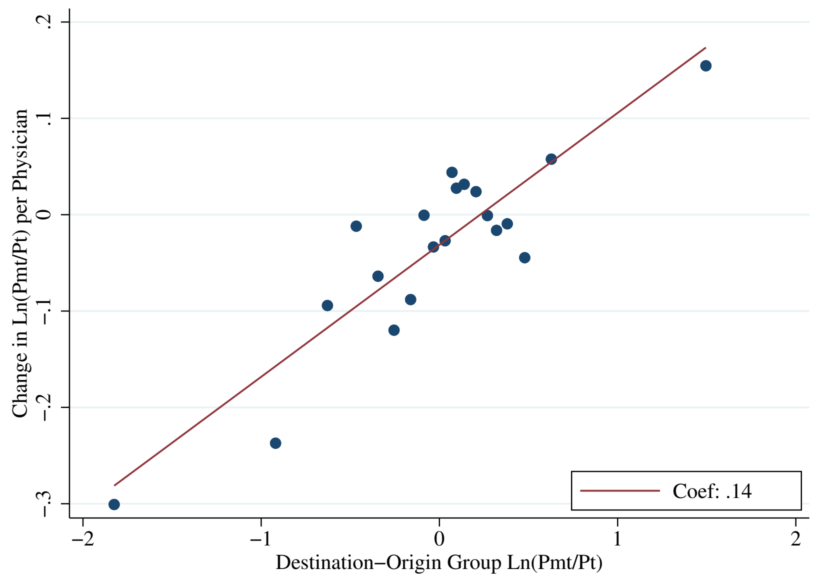
(b) $\Delta_{pmt/pt} > 0$



(c) $\Delta_{pmt/pt} < 0$

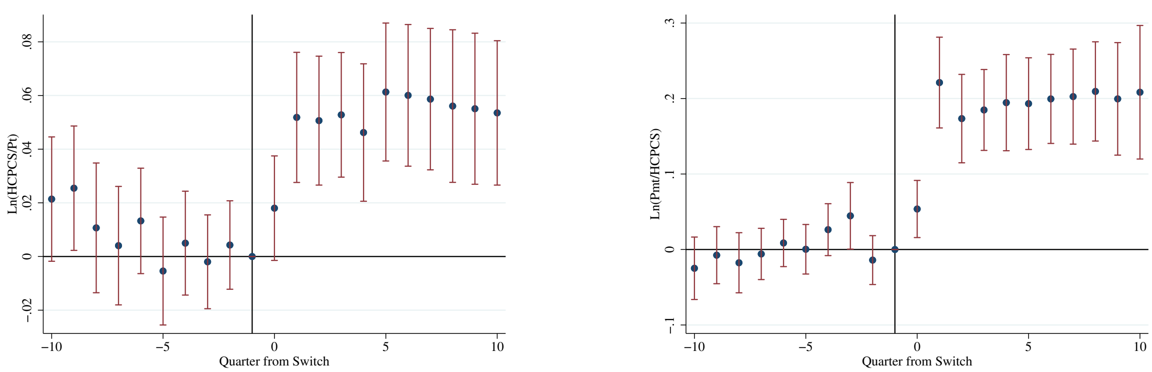
This figure plots the θ_q's from Equation 6, estimated for log reimbursement per patient-quarter at the physician level. In panels (b) and (c), we estimate Equation 6 separately for $\Delta_{pmt/pt} > 0$ and $\Delta_{pmt/pt} < 0$, respectively. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

Figure 3: Changes in Physician Treatment Intensity v. Changes in Group Intensity



This figure plots the vigintiles of the change in group intensity (as captured by log reimbursement per patient-quarter per physician; x-axis) and corresponding average change in physician intensity (similarly captured by log reimbursement per patient-quarter; y-axis). The line of best fit is given by a simple OLS regression of the 20 data points associated with the change in physician intensity on the change in group intensity. "Coef" is the slope of the line through these points.

Figure 4: Additional Physician Treatment Intensity Measures Relative to a Change in Group Intensity

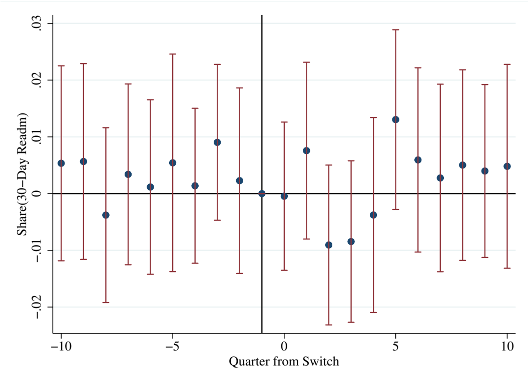


(a) Ln(HCPCS/Patient)

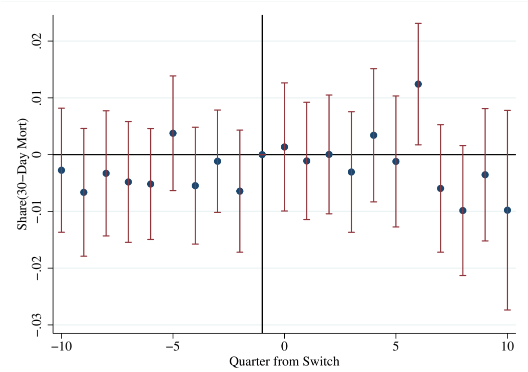
(b) Ln(Payment/HCPCS)

This figure plots the θ_q's estimated from Equation 6 for additional measures of treatment intensity. HCPCS is the Healthcare Common Procedural Coding System code recorded as the specific line item in a given claim as a measure of the quantity of claims; Payment/HCPCS measures the average payment per claim item. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

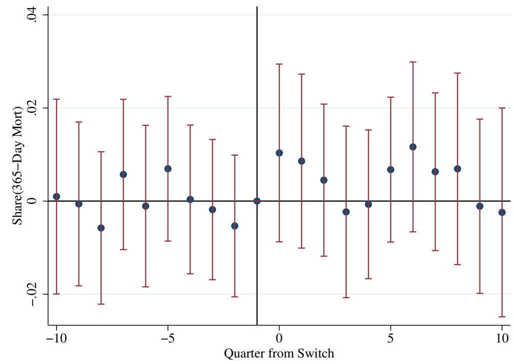
Figure 5: Physician Quality of Care Relative to a Change in Group Intensity



(a) 30-Day Readmission



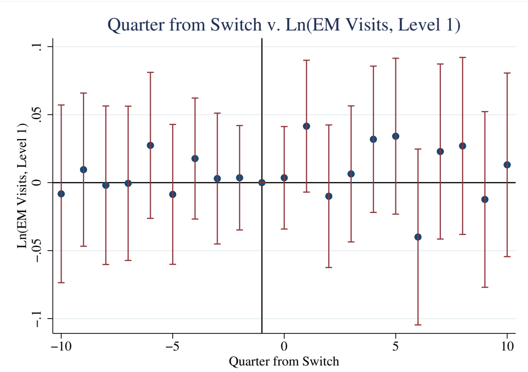
(b) 30-Day Mortality



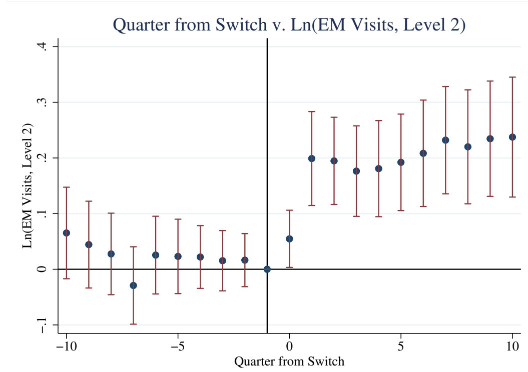
(c) One-Year Mortality

This figure plots the θ_q's estimated from Equation 6 for measures of quality of care. 30-day readmission rates are calculated as the share of hospitalizations in a given quarter that resulted in a readmission within 30 days of the discharge date. 30- and 365-day mortality rates are calculated as the share of hospitalizations in which the patient died within 30 or 365 days of admission. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

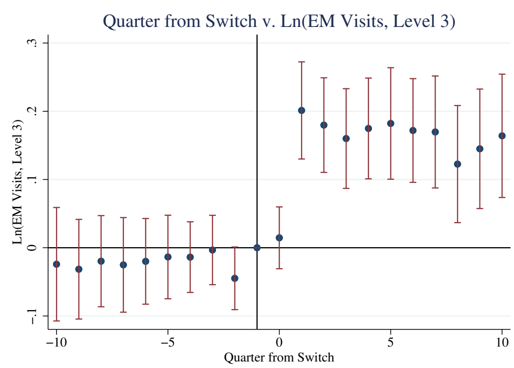
Figure 6: E&M Visits by Intensity Level Relative to a Change in Group Intensity



(a) Level 1



(b) Level 2



(c) Level 3

This figure plots the θ_qs estimated from Equation 6, scaled by $\Delta_{pmt/pt}$. The outcomes are the log number of Evaluation and Management (E&M) visits of a particular level of intensity (levels 1 through 3) as described in the text. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

Appendix For Online Publication:

Physician Group Influences on Treatment Intensity and Health:
Evidence from Physician Switchers

Joseph J. Doyle (jjdoyle@mit.edu) and Becky Staiger (bstaiger@berkeley.edu)

Appendix A Additional Tables and Figures

Table A.1: Physician Count by Restriction

Restriction	Count
Physicians with a claim that can be linked to a beneficiary's inpatient stay	553721
Physicians with 90% of claims in a given year-quarter associated with one group	552420
Switchers	
Physicians who belong to an origin group for at least four consecutive quarters, and then switch to a destination group where they are subsequently observed for at least five consecutive quarters (including the switch quarter)	72426
Physicians who remain in one hospital throughout their episode	30488
Physicians whose origin and destination groups exist in the four-quarter pre-switch period	19847
Physicians who are in groups with at least one other physician	16187
Physicians in origin and destination groups that treat at least 10 patients per quarter	13883
With a specialty of internal medicine (Internists)	3108
Non-Switchers	
Physicians who are only ever in one group	321963
Physicians who remain in one hospital throughout their episode	237496
Physicians in hospitals with switcher physicians	162433
With a specialty of internal medicine (Internists)	30887

This table reports the number of physicians at each step of the sample construction, after imposing a particular restriction.

Table A.2: Difference-in-Differences, Additional Specialties

	(1) Ln(Pmt/Pt)		(2) Ln(Pmt)		(3) Share(30-Day Readm)		(4) Share(30-Day Mort)		(5) Share(365-Day Mort)	
Panel A. All Specialties										
$\Delta_{pmt/pt}$*Post Switch	0.085***	(0.009)	0.052***	(0.011)	0.001	(0.001)	-0.001	(0.001)	-0.001	(0.001)
$\Delta_{pmt/pt}$*Qtr=0	0.011	(0.007)	-0.024**	(0.009)	0.002	(0.002)	-0.002	(0.001)	-0.002	(0.002)
Constant	5.193***	(0.000)	7.083***	(0.000)	0.246***	(0.000)	0.093***	(0.000)	0.282***	(0.000)
Observations	2997461		2997461		2997461		2997461		2997461	
Dep. Var. Mean	5.192		7.082		0.246		0.093		0.282	
Panel A. Not Internal Medicine										
$\Delta_{pmt/pt}$*Post Switch	0.060***	(0.007)	0.021*	(0.008)	0.002	(0.001)	-0.000	(0.001)	-0.001	(0.001)
$\Delta_{pmt/pt}$*Qtr=0	0.005	(0.007)	-0.031**	(0.009)	0.002	(0.002)	-0.002	(0.002)	-0.004	(0.002)
Constant	5.176***	(0.000)	7.047***	(0.000)	0.246***	(0.000)	0.091***	(0.000)	0.276***	(0.000)
Observations	2450115		2450115		2450115		2450115		2450115	
Dep. Var. Mean	5.175		7.045		0.246		0.091		0.276	

Fixed effects for physician-episode and hospital-year-quarter. Standard errors in parentheses are two-way clustered at the physician and group levels. Omitted category is an indicator for quarters $\in [-10, -1]$.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.3: Compare Specifications

	(1) Ln(Pmt/Pt)	(2) Ln(Pmt/Pt)	(3) Ln(Pmt/Pt)
$\Delta_{pmt/pt}$*Post Switch	0.266***	0.266***	0.256***
	(0.037)	(0.038)	(0.039)
$\Delta_{pmt/pt}$*Qtr=0	0.067**	0.067**	0.064**
	(0.023)	(0.023)	(0.024)
Constant	5.268***	5.268***	5.269***
	(0.001)	(0.001)	(0.001)
Observations	529465	529465	528154
Adjusted R^2	0.674	0.674	0.674
Dep. Var. Mean	5.267	5.267	5.268
Specification	Main	Bayes	Non-Switchers

Fixed effects for physician-episode and hospital-year-quarter. Standard errors in parentheses are two-way clustered at the physician and group levels. Omitted category is an indicator for quarters $\in [-10, -1]$. Column (1) re-states the main results. Column (2) reports the results estimated using the empirical Bayes adjusted measure of $\Delta_{pmt/pt}$. Column (3) reports results estimated for a version of $\Delta_{pmt/pt}$ that is calculated only based on non-switching physicians in the origin and destination groups.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.4: Difference-in-Differences by Hospital Medicare Share

	(1) Ln(Pmt/Pt)	(2) Ln(Pmt/Pt)	(3) Ln(Pmt/Pt)	(4) Ln(Pmt/Pt)
$\Delta_{pmt/pt}$*Post Switch	0.376***	0.181**	0.222***	0.242***
	(0.079)	(0.056)	(0.048)	(0.053)
$\Delta_{pmt/pt}$*Qtr=0	0.084	-0.018	0.086**	0.087*
	(0.049)	(0.050)	(0.031)	(0.042)
Constant	5.269***	5.268***	5.268***	5.268***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	529465	529465	529465	529465
Dep. Var. Mean	5.272	5.274	5.272	5.272
Medicare Share Quantile	1	2	3	4
Share Range	.045–.396	.397–.457	.457–.509	.509–1.231

Estimates come from a single model with interactions for the different Medicare-share quartiles, including fixed effects for physician-episode and hospital-year-quarter. Standard errors in parentheses are two-way clustered at the physician and group levels. Omitted category is an indicator for quarters $\in [-10, -1]$. Medicare share quantiles are calculated for a physician's attributed hospital in a given year-quarter and are based on the American Hospital Association annual survey. Shares for a small number of hospitals exceed one due to measurement error.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: Difference-in-Differences, Additional Outcomes

	(1) Ln(Patients)		(2) Ln(Pmt)		(3) Ln(HCPCS)		(4) Ln(Procedures)		(5) Ln(LOS)	
Panel A. $\Delta_{pmt/pt}$										
$\Delta_{pmt/pt}$*Post Switch	0.010	(0.023)	0.290***	(0.056)	0.078**	(0.030)	0.022	(0.022)	0.001	(0.006)
$\Delta_{pmt/pt}$*Qtr=0	-0.027	(0.016)	0.036	(0.032)	-0.010	(0.020)	-0.013	(0.016)	-0.003	(0.014)
Constant	2.184***	(0.001)	7.250***	(0.001)	3.058***	(0.001)	1.705***	(0.001)	2.046***	(0.000)
Dep. Var. Mean	2.185		7.250		3.059		1.706		2.046	
Panel B. Δ_{size}										
Δ_{size}*Post Switch	-0.030***	(0.008)	-0.050***	(0.012)	-0.042***	(0.010)	-0.026***	(0.008)	-0.003	(0.003)
Δ_{size}*Qtr=0	-0.028***	(0.007)	-0.026*	(0.012)	-0.026**	(0.009)	-0.022**	(0.007)	0.003	(0.005)
Constant	2.184***	(0.001)	7.250***	(0.001)	3.058***	(0.001)	1.704***	(0.001)	2.046***	(0.000)
Dep. Var. Mean	2.185		7.250		3.059		1.706		2.046	
Observations	529465		529465		529465		529465		529465	

This table reports estimated coefficients from Equation 7 for internists. Post Switch is an indicator variable that is equal to 1 for all quarters $\in [1, 10]$. Fixed effects are included for physician-episode and hospital-year-quarter. Standard errors in parentheses are two-way clustered at the physician and group levels. Omitted category is an indicator for quarters $\in [-10, -1]$. HCPCS is the Healthcare Common Procedural Coding System code recorded as the specific line item in a given claim. "Pt" abbreviates patient. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6: Inpatient Evaluation and Management Codes, By Level

Level	CPT Code	Description	Avg. Reimbursement
1	99221	Hospital initial inpatient care, straightforward or low complexity	\$97.40
	99231	Subsequent inpatient care, straightforward or low complexity	\$40.64
	99234	Admission and discharge same day, straightforward or low complexity	\$130.80
2	99222	Hospital initial inpatient care, moderate complexity	\$132.44
	99232	Subsequent inpatient care, moderate complexity	\$74.24
	99235	Admission and discharge same day, moderate complexity	\$166.79
3	99223	Hospital initial inpatient care, high complexity	\$194.89
	99233	Subsequent inpatient care, high complexity	\$104.69
	99236	Admission and discharge same day, high complexity	\$211.83

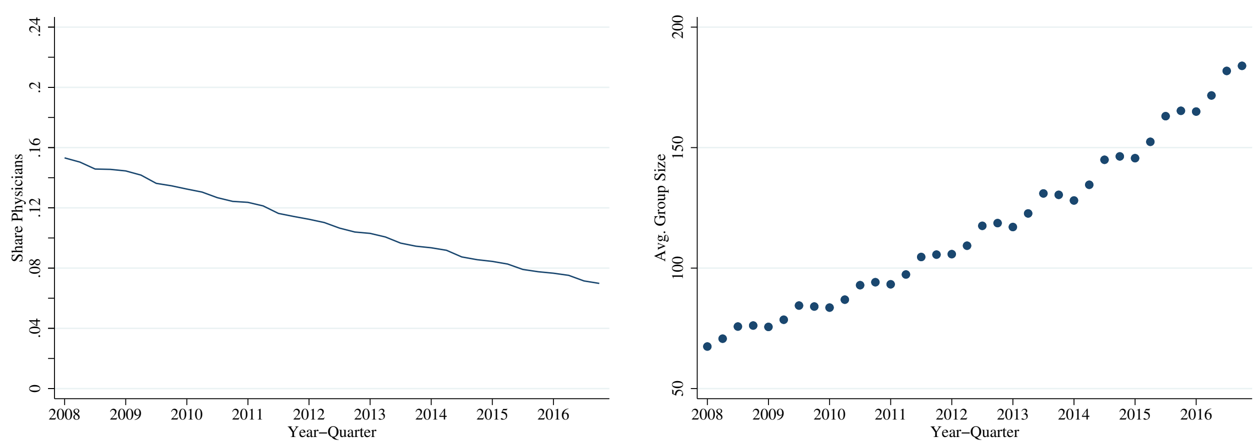
This table reports the inpatient evaluation and management (E&M) codes used in our analysis of billing intensity. Common inpatient E&M codes were identified from University of Southern California medical group compliance guidelines. Average reimbursement is calculated from the carrier files as the sum of the line NCH payment amount, the line beneficiary part B deductible amount, the line coinsurance amount, and the line beneficiary primary payer paid amount. The resulting total represents the payment due to the provider for that particular HCPCS.

Table A.7: Horse Race Model of Change in Intensity and Change in Size

	(1) Ln(Pmt/Pt)	(2) Ln(Pmt/Pt)	(3) Ln(Pmt/Pt)
$\Delta_{pmt/pt}$*Post Switch	0.266***		0.268***
	(0.037)		(0.037)
$\Delta_{pmt/pt}$*Qtr=0	0.067**		0.067**
	(0.023)		(0.023)
Δ_{size}*Post Switch		-0.015	-0.020**
		(0.008)	(0.007)
Δ_{size}*Qtr=0		0.005	0.005
		(0.007)	(0.007)
Constant	5.268***	5.268***	5.268***
	(0.001)	(0.001)	(0.001)
Observations	529465	529465	529465
Dep. Var. Mean	5.267	5.267	5.267

All models include fixed effects for physician-episode and hospital-year-quarter. Standard errors in parentheses are two-way clustered. Δ_{size} is the change in group size as measured by the number of physicians.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure A.1: Group Trends in Hospitals Over Time

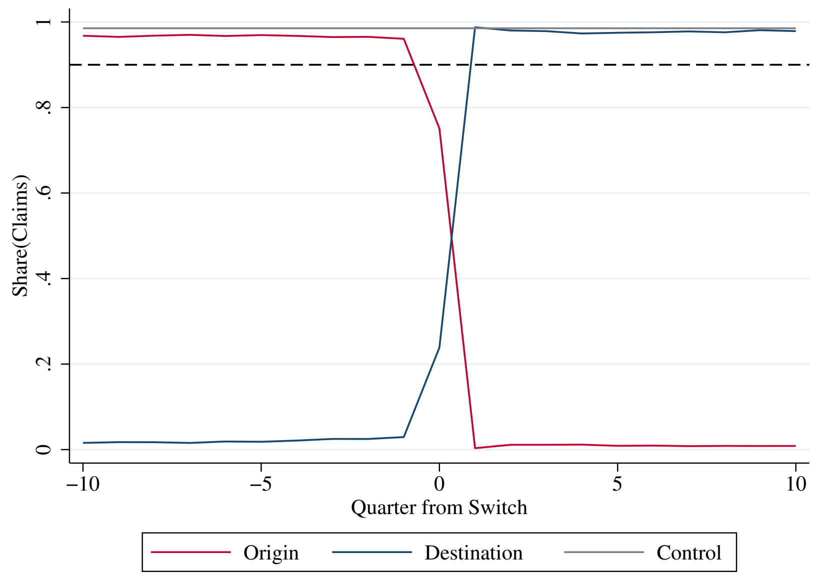


(a) Share(Physicians) In Solo Practice

(b) Group Size

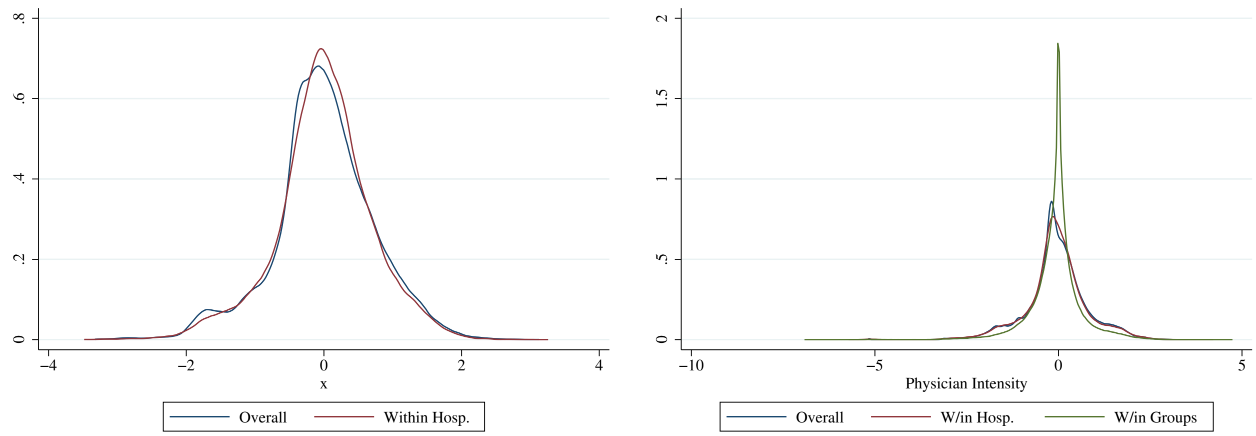
This figure plots trends in group size among physicians practicing in hospitals by year-quarter during our study period (2008-2016). Panel (a) plots the share of physicians in a group size of 1 (which we characterize as solo practice) over time. Panel (b) plots average group size over time. Group size is calculated as the number of distinct National Provider Identifiers (NPIs) with an entity type of "1" (i.e. an individual) and a taxonomy type of "Allopathic & Osteopathic Physicians" associated with a particular billing identifier.

Figure A.2: Share of HCPCS Associated with a Group, Internists



This figure plots the average share of physicians' HCPCS that are associated with their origin and destination groups in the quarters relative to the switch. The red line plots the share of HCPCS associated with the origin group of the switching physician in a given quarter relative to the switch; the blue line, the share associated with the destination group of the switching physician. The gray line plots the average share of HCPCS associated with the single group that a non-switching physician belongs to over all quarters (by definition, non-switchers don't have quarters relative to a switch). The dotted black line indicates the 0.9 threshold (generally) used to attribute physicians to groups.

Figure A.3: Distribution of Group and Physician Intensity

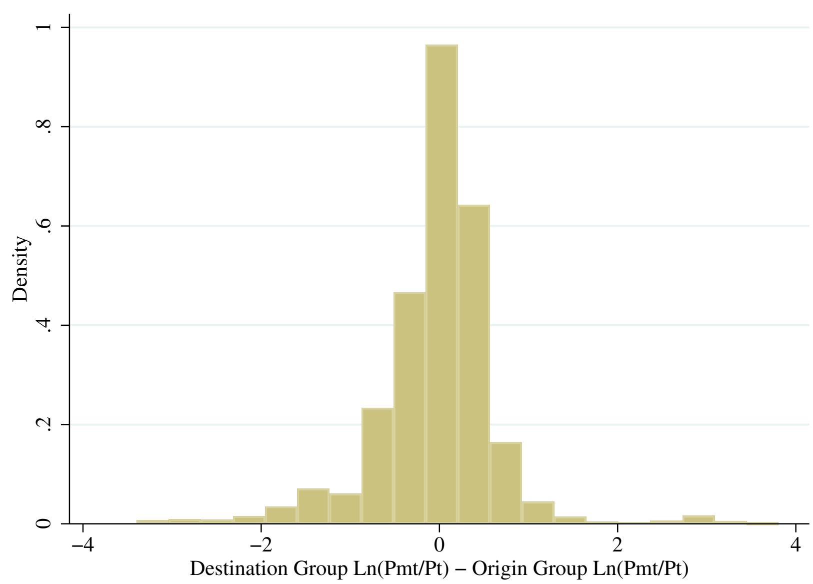


(a) Group Intensity, All Specialties

(b) Physician Intensity, All Specialties

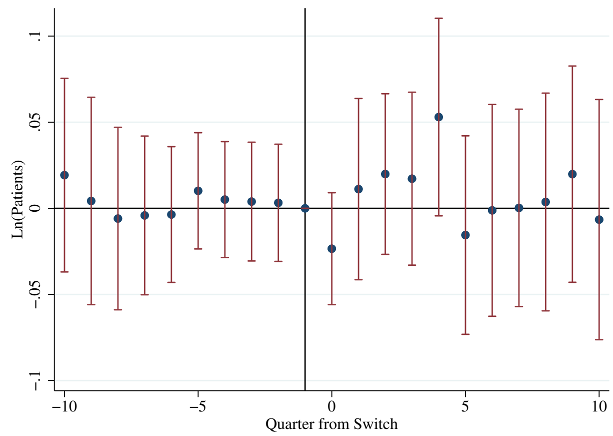
This figure plots variation in physician and group intensity among physicians of all specialties, as measured by average log reimbursement per patient-quarter per physician. Group intensity is calculated as described above, and average physician intensity is calculated across all quarters. Panels (a) and (b) plot the variation in (demeaned) physician and group intensity overall, within hospitals, and within groups (for physicians only), for switchers, non-switchers, and all other out-of-sample physicians associated with in-sample groups. Within-hospital and within-group intensity is demeaned using the hospital- and hospital-group specific averages, respectively. The standard deviation for overall and within-hospital group intensity is 0.73 and 0.70, respectively. The standard deviation for overall, within-hospital, and within-group intensity for physicians is 0.84, 0.84, and 0.66, respectively.

Figure A.4: Distribution of $\Delta_{pmt/pt}$

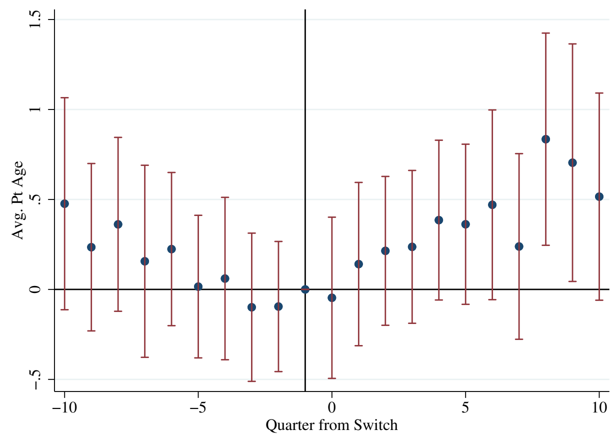


This figure plots the distribution of (non-demeaned) $\Delta_{pmt/pt}$ for physicians in our main empirical sample. $\Delta_{pmt/pt}$ has a mean of -0.032 and a standard deviation of 0.68.

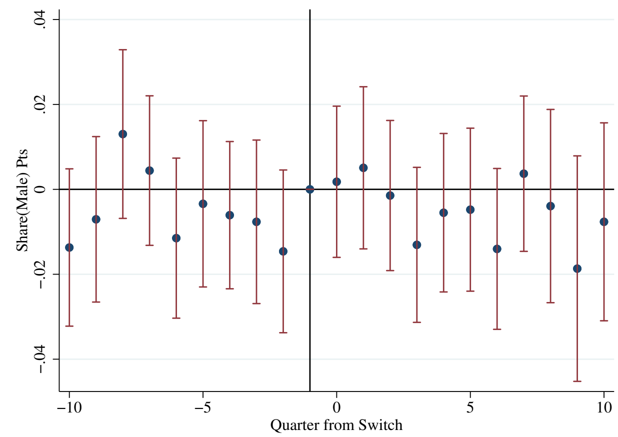
Figure A.5: Test for Balance of Patient Characteristics Across Switch, Internists



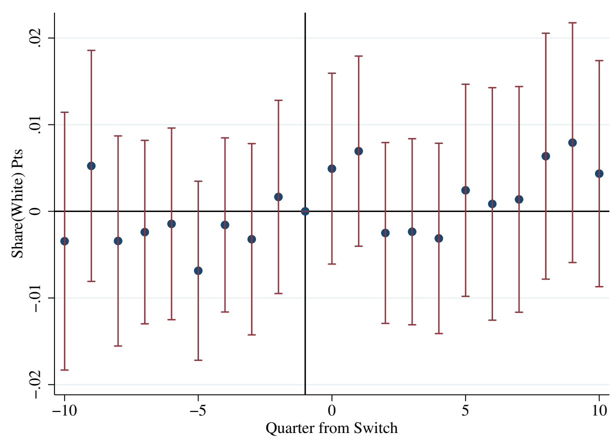
(a) Log Patients Across Switch



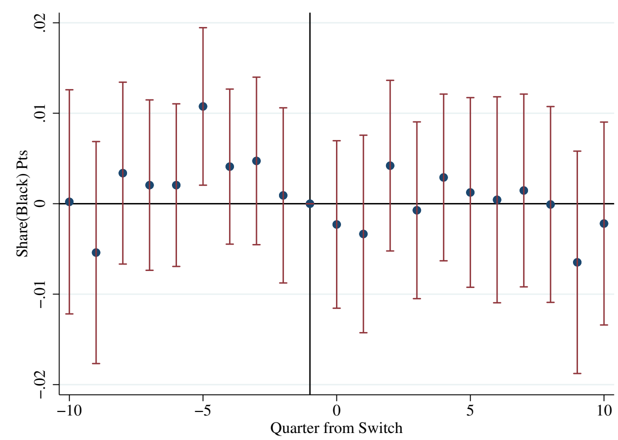
(b) Patient Age Across Switch



(c) Patient Sex Across Switch



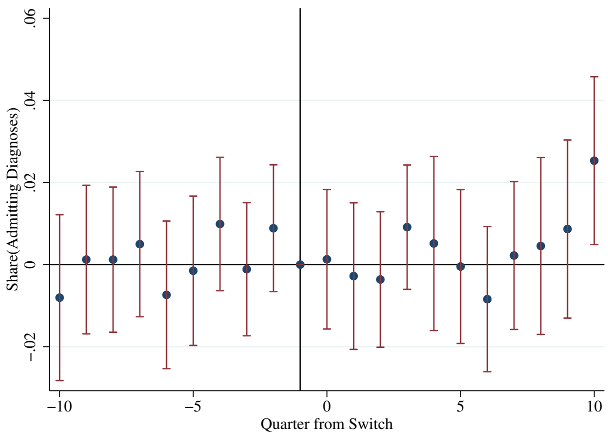
(d) Share of White Patients Across Switch



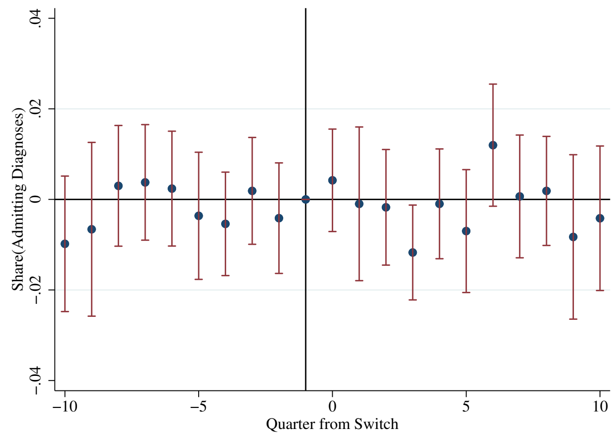
(e) Share of Black Patients Across Switch

This figure plots changes in patient characteristics across the switch, scaled by $\Delta_{pmt/pt}$ (i.e. plots of θ_qs from Equation 6 with patient characteristics as the dependent variables). Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

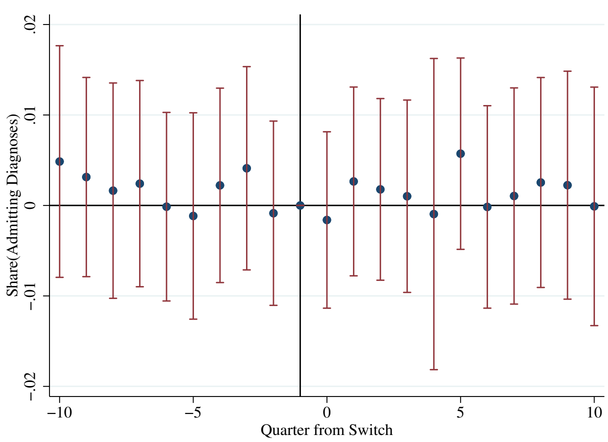
Figure A.6: Most Common ICD Admitting Diagnostic Sections, 1-5



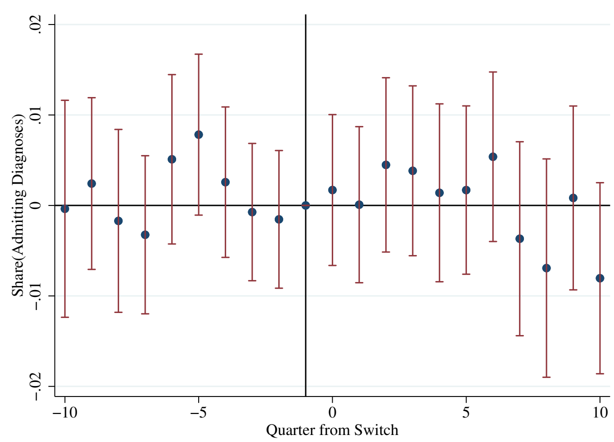
(a) Circulatory System



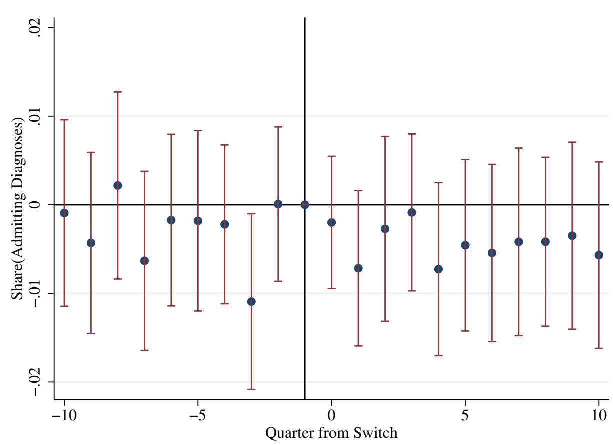
(b) Symptoms, Signs, and Ill-Defined Conditions



(c) Respiratory System



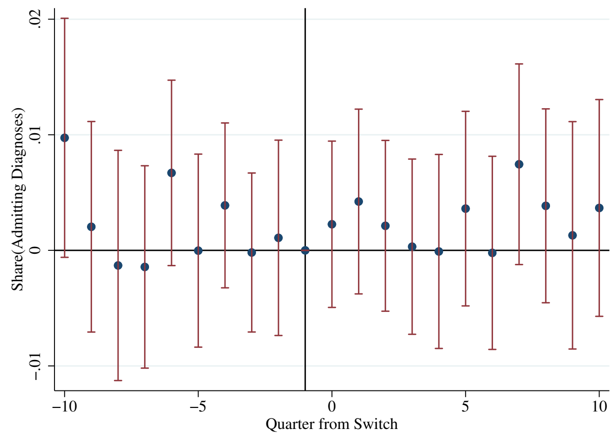
(d) Genitourinary System



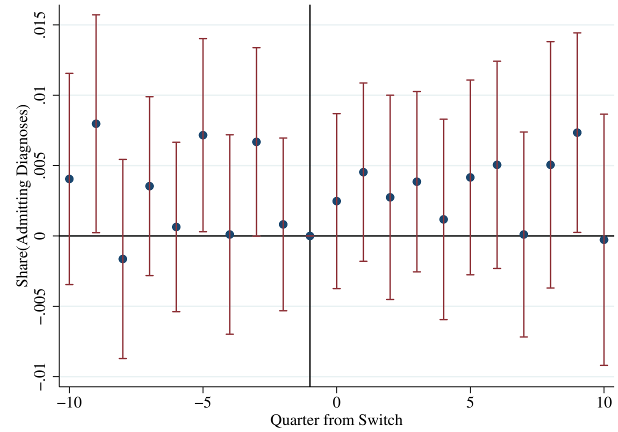
(e) Digestive System

This figure plots changes in the shares of the top five most common hierarchical ICD-CM-9 and ICD-CM-10 sections across a physician's switch between groups, scaled by $\Delta_{pmt/pt}$ (i.e. plots of θ_qs from Equation 6 with ICD sections as the dependent variable). Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

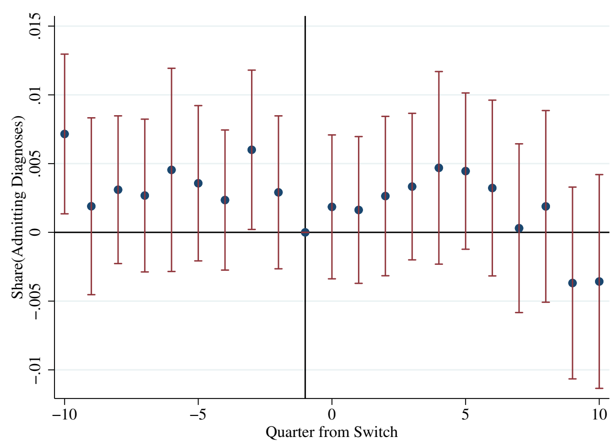
Figure A.7: Most Common ICD Admitting Diagnostic Sections, 6-10



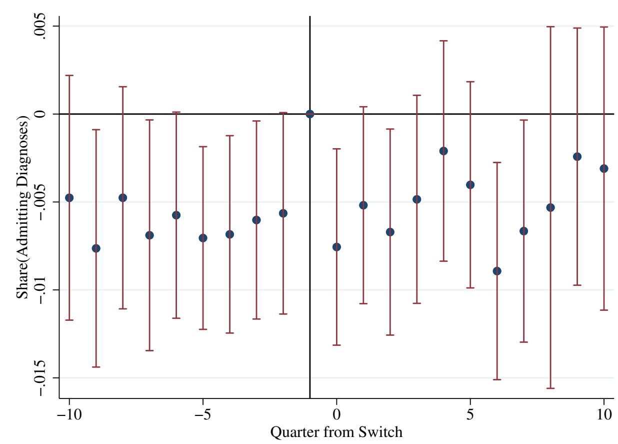
(a) Endocrine, Nutritional and Metabolic Diseases, and Immunity Disorders



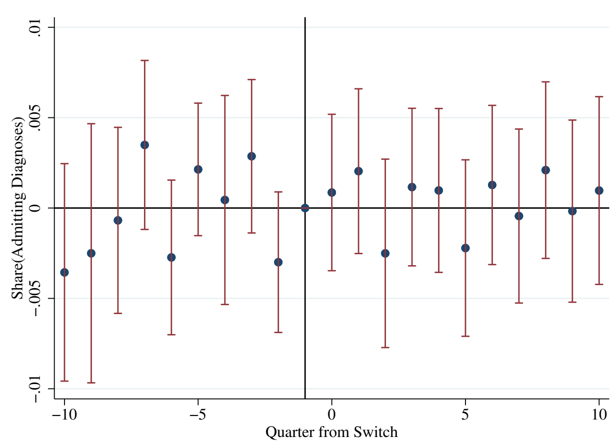
(b) Injury and Poisoning



(c) Infectious and Parasitic Diseases



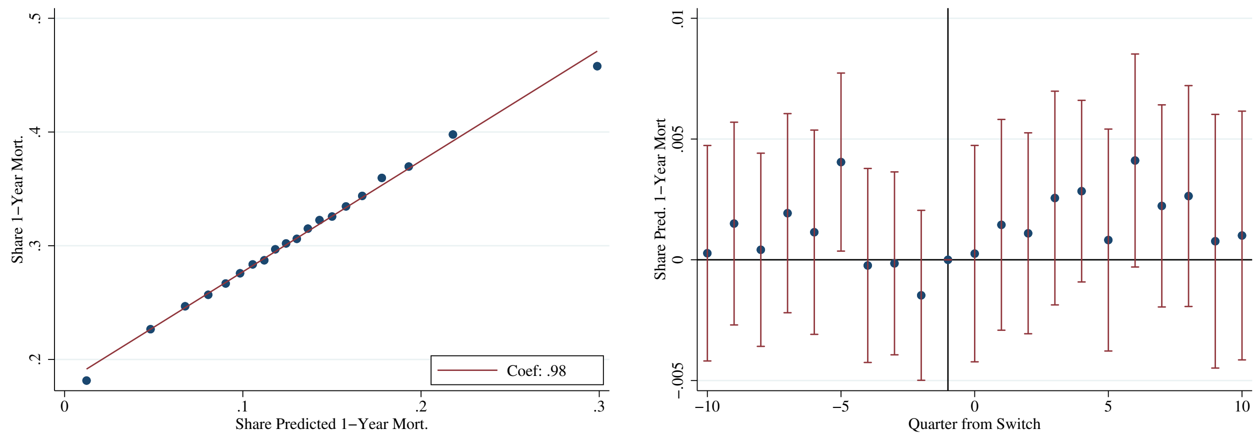
(d) Musculoskeletal System and Connective Tissue



(e) Blood and Blood-forming Organs

This figure plots changes in the shares of the top six through tenth most common hierarchical ICD-CM-9 and ICD-CM-10 sections across a physician's switch between groups, scaled by $\Delta_{pmt/pt}$ (i.e. plots of θ_qs from Equation 6 with ICD sections as the dependent variable). Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

Figure A.8: Balance in Predicted Mortality

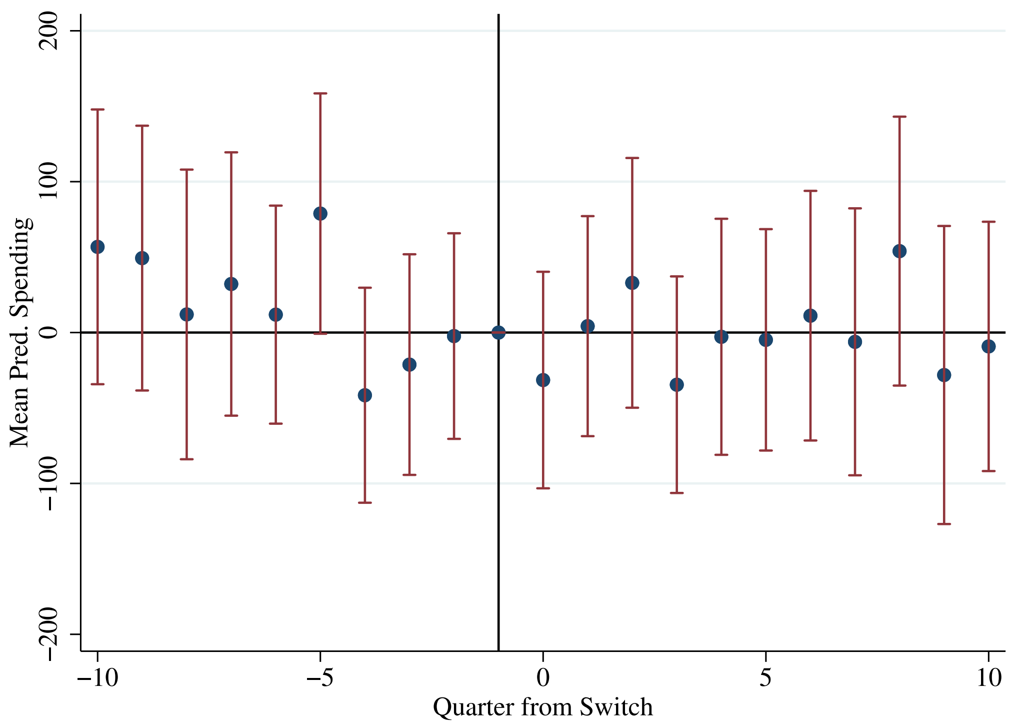


(a) Share(1-Year Mort.) v. Share(Predicted 1-Year Mort.)

(b) Average Patient Predicted 1-Year Mortality Across Switch

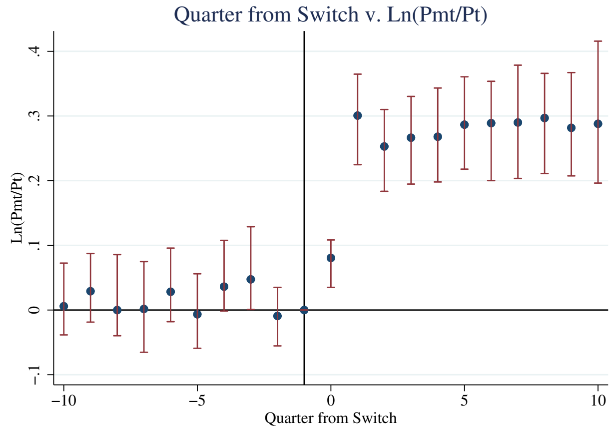
This figure shows balance in predicted patient mortality across the group switch. Panel (a) plots the relationship between the observed share of a physician's patients with 1-year mortality and the share of a physician's patients with predicted 1-year mortality (based on the approach discussed below). We find the vigintiles of the share of predicted mortality, and collapse the observations (at the physician-quarter level) of the observed share and predicted share to their means, plotted here. The coefficient of 0.98 represents the relationship between the share of predicted 1-year mortality and the observed share of 1-year mortality within these vigintile bins. Panel (b) plots changes in average predicted 1-year mortality of patients across the switch, scaled by $\Delta_{pmt/pt}$ (i.e. plots of θ_qs from Equation 6 with predicted mortality as the dependent variable.) Predicted 1-year mortality is calculated in the following steps. First, we estimate the relationship (in a linear model) between an indicator for whether a patient died in 2012 or 2013 and patient age (in vigintiles), sex, race, and comorbidity indicators recorded in 2012, with 2012 being the midpoint of our study period. We exclude all patients treated by physicians in our final study sample from this analysis. Using the coefficients obtained from this regression, we predict 1-year mortality for each patient treated by a physician in our final study sample based on the inputs in the model. Notably, to avoid endogeneity concerns of a new group's influence on diagnostic intensity, we use the comorbidity indicators from the year prior to the physician treating the patient (i.e. the index treatment event). We collapse predicted one-year mortality to the physician-quarter level, and estimate our main event study for this outcome. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

Figure A.9: Predicted Spending



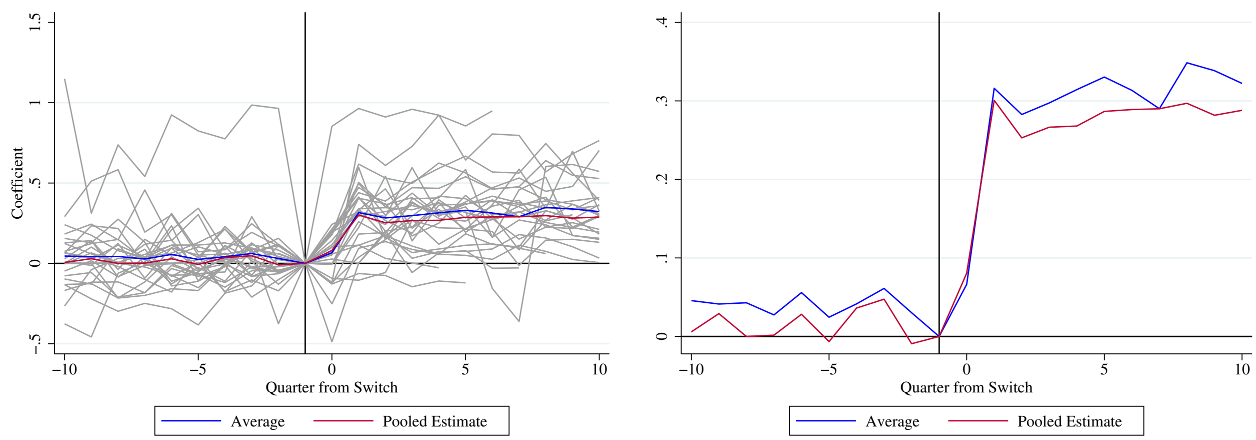
Predicted inpatient spending is calculated in the following steps. First, we estimate the relationship (in a linear model) between inpatient spending associated with a given hospitalization and patient age (in vigintiles), sex, race, and comorbidity indicators recorded in 2012, with 2012 being the midpoint of our study period. We exclude all patients treated by physicians in our final study sample from this analysis. Using the coefficients obtained from this regression, we predict inpatient spending for each patient treated by a physician in our final study sample based on the inputs in the model. Notably, to avoid endogeneity concerns of a new group's influence on diagnostic intensity, we use the comorbidity indicators from the year prior to the physician treating the patient (i.e. the index treatment event). We collapse predicted inpatient spending to the physician-quarter level, and estimate our main event study for this outcome. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

Figure A.10: Physician Treatment Intensity Relative to a Change in Group Intensity, Bootstrapped 95% Confidence Intervals



This graph plots the main results with bootstrapped 95% confidence intervals to take into account variability due to the calculation of $\Delta_{pmt/pt}$ as a generated regressor. To do this, we first empirically re-sample (with replacement) the data at the claims level for each origin and destination group, iterating 50 times, to generate a set of simulated origin and destination group intensities which we then use to calculate a set of (50) $\Delta_{pmt/pt}$s for each switching physician. Next, we re-sample each simulated dataset (with replacement) at the switching physician-episode level, and re-run Equation 6 50 times to estimate our 95% confidence interval using the coefficients estimated from these 50 iterations. Due to computational limitations, we only show the bootstrapped 95% confidence intervals for our main result to illustrate that this approach does not change the significance of our main findings.

Figure A.11: Treatment Intensity Event Study by Switch-Quarter Cohort, Internists

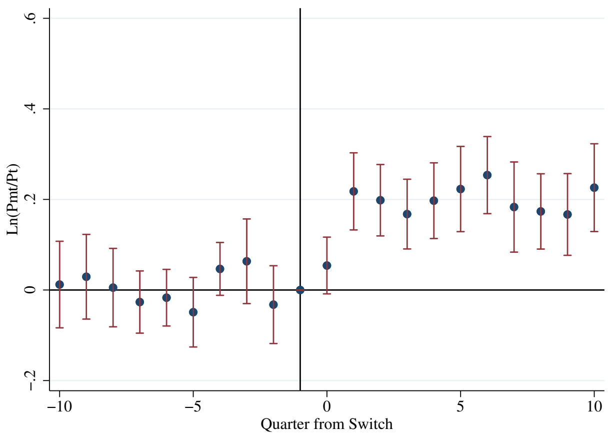


(a) With Switcher Cohorts

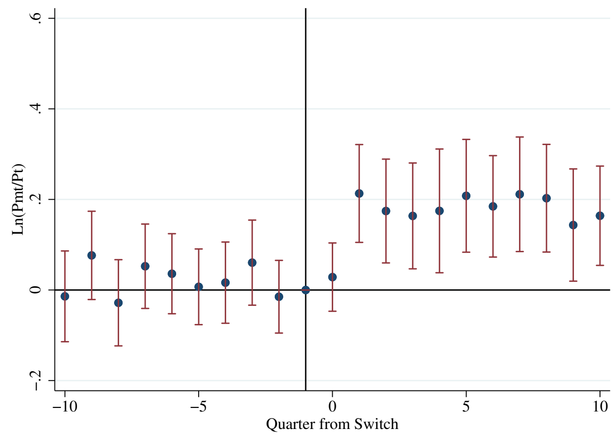
(b) Average v. Pooled Estimates

This figure plots the θ_qs obtained from estimating Equation 6 on each cohort of switchers, as defined by quarter of switch. Estimates of difference in intensity relative to switch time (relative to controls) for individual switcher cohorts are plotted in gray in (a). The blue line plots the average of these estimates, and the red line plots the original estimates from running the model on the pooled (all switcher cohort) sample in (a) and (b). 95% confidence intervals are excluded for ease of comparison of the different trend lines.

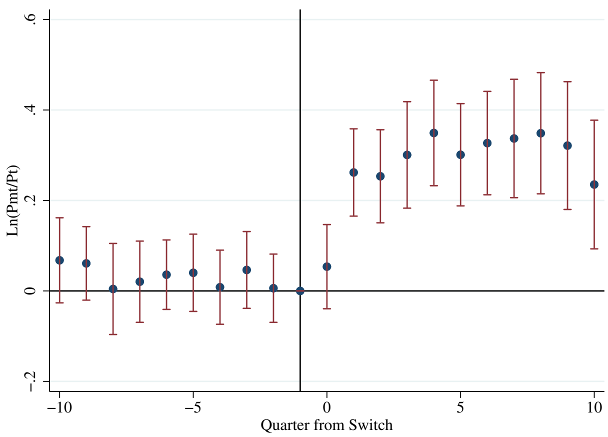
Figure A.12: Physician Treatment Intensity Relative to a Change in Group Intensity, By Quartile of Physicians' Years Experience



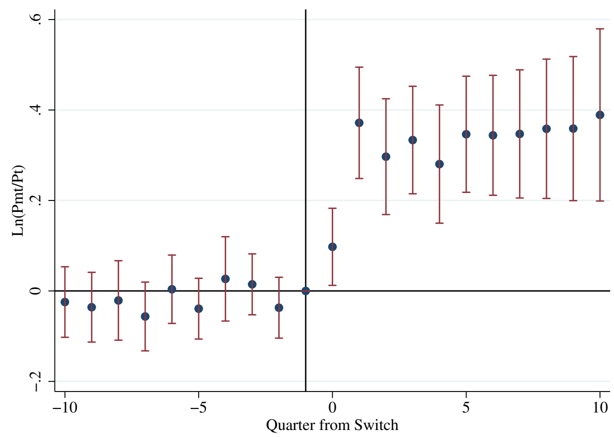
(a) Years Experience: 2 - 14



(b) Years Experience: 15 - 19



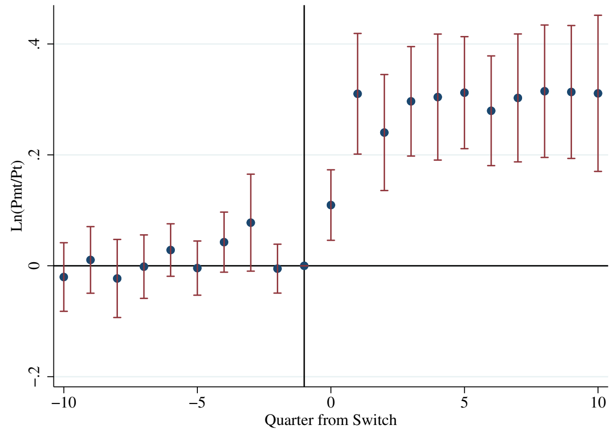
(c) Years Experience: 20 - 27



(d) Years Experience: 28 - 56

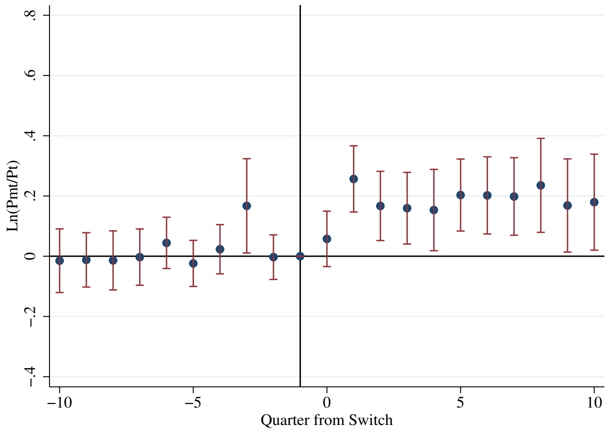
This figure plots the θ_q's estimated from Equation 6, scaled by $\Delta_{pmt/pt}$, by quartiles of years of experience, as measured by years from medical school graduation relative to 2016. Quartiles are defined by switchers' years of experience. Year of graduation is obtained from the Physician Compare database. Of the 3,242 internist physician-episodes in the sample, 2,986 (92%) can be matched to the Physician Compare database. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

Figure A.13: Physician Treatment Intensity Relative to a Change in Group Intensity, Scaled by Destination Intensity

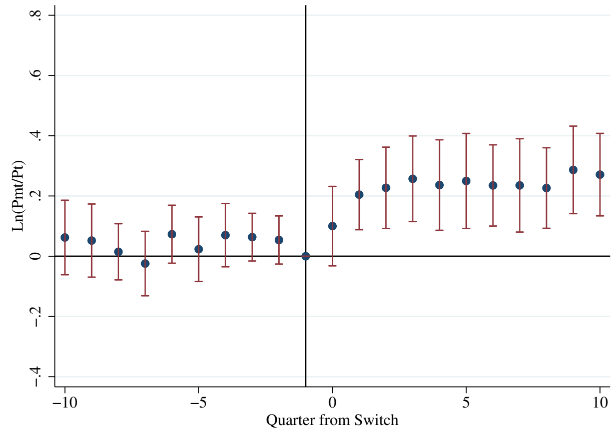


This figure plots the θ_qs when scaled by the destination group's pre-switch intensity instead of $\Delta_{pmt/pt}$. The coefficients around the switch can be interpreted as the increase in physician intensity corresponding to an increase in (pre-switch) destination group intensity before physician p joins that group. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

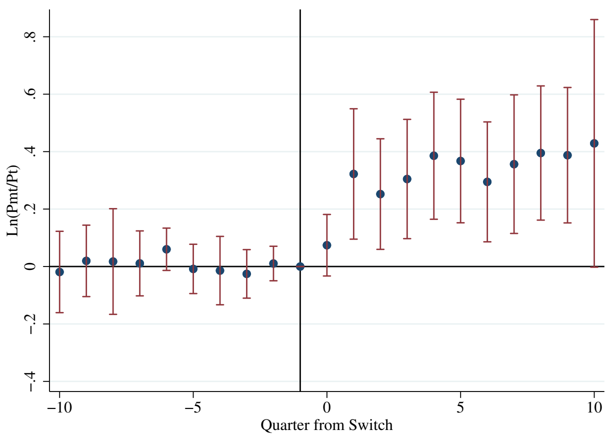
Figure A.14: Physician Treatment Intensity Relative to a Change in Group Intensity, by Quartile of Origin Group Intensity



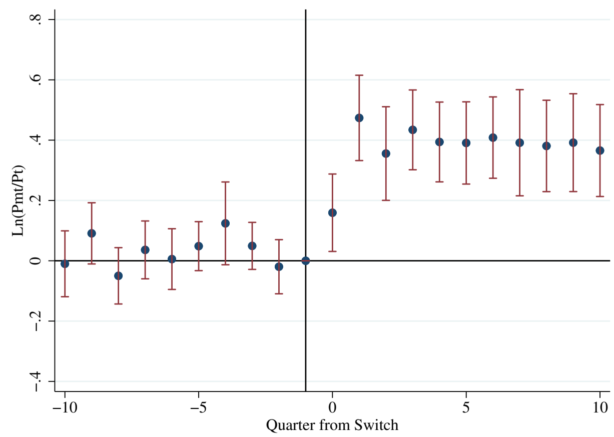
(a) Origin Intensity, 2.04 - 5.2



(b) Origin Intensity, 5.2 - 5.35



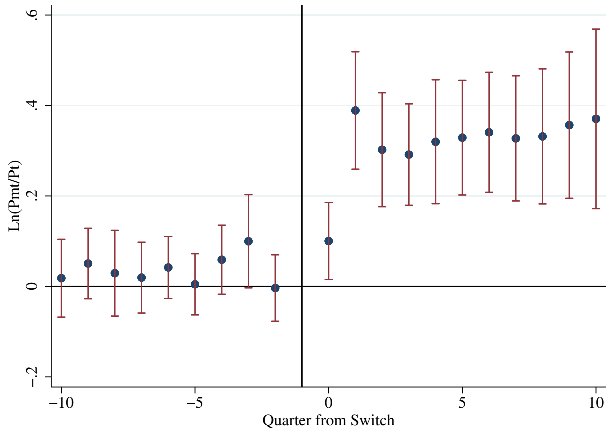
(c) Origin Intensity, 5.35 - 5.52



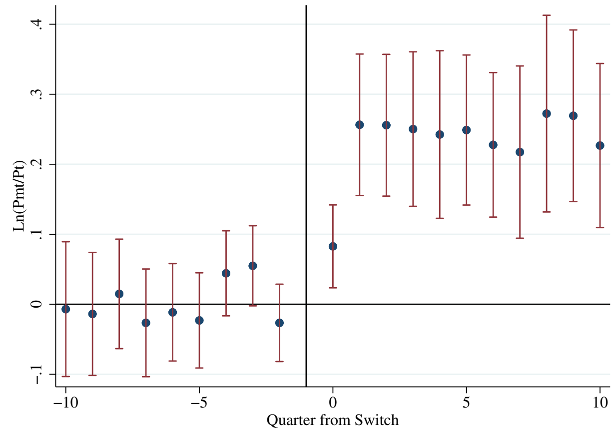
(d) Origin Intensity, 5.52 - 6.56

This figure plots the θ_qs estimated from Equation 6 by quartile of origin group intensity, as measured by average pre-switch log reimbursement per physician per quarter (excluding the switching physician) in $q \in [-4, -1]$. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

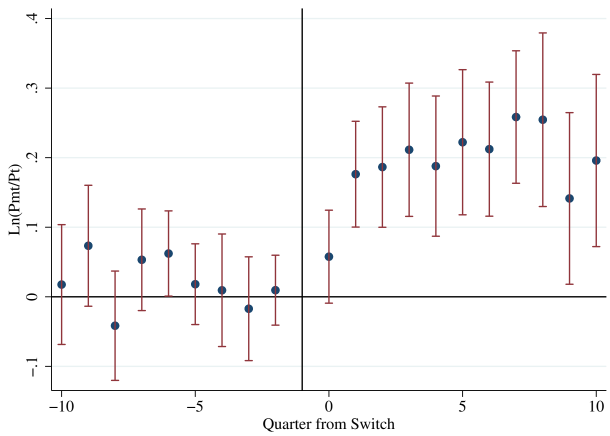
Figure A.15: Physician Treatment Intensity Relative to a Change in Group Intensity, by Quartile of Physician's Pre-Switch Intensity



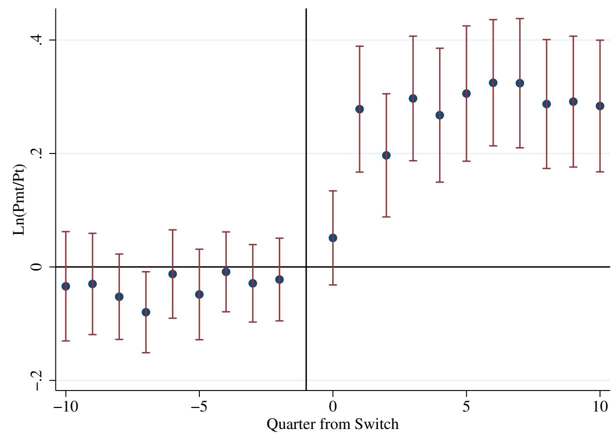
(a) Ln(Pmt/pt): 3.61 - 7.36



(b) Ln(Pmt/pt): 7.36 - 7.92



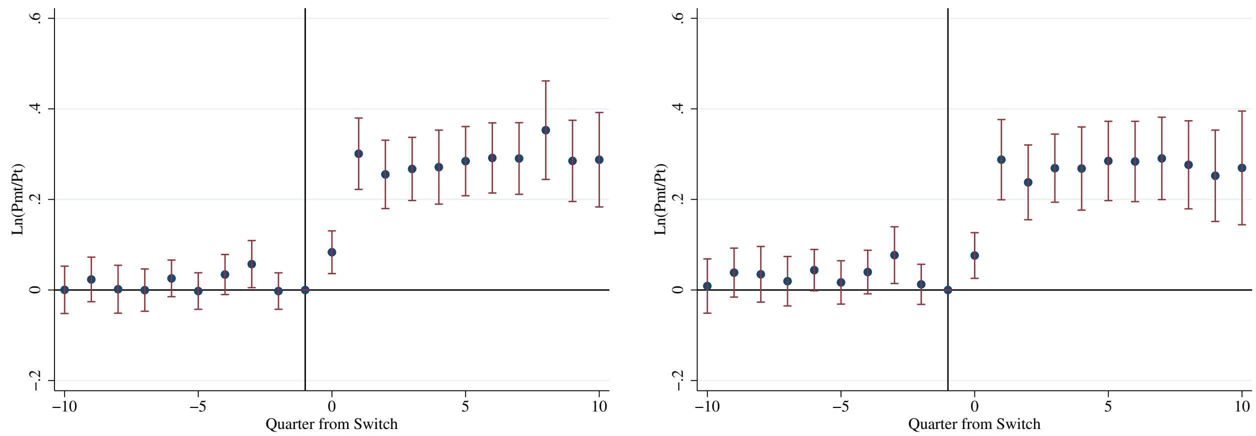
(c) Ln(Pmt/pt): 7.93 - 8.33



(d) Ln(Pmt/pt): 8.34 - 9.9

This figure plots the θ_qs estimated from Equation 6 by quartile of a physician's pre-switch intensity. Pre-switch intensity is measured as the average ln(reimbursement/patient) per physician-quarter across pre-switch quarters $q \in [-10, -1]$. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

Figure A.16: Physician Treatment Intensity Relative to a Change in Group Intensity, 5 and 20 Minimum Patients per Group per Quarter

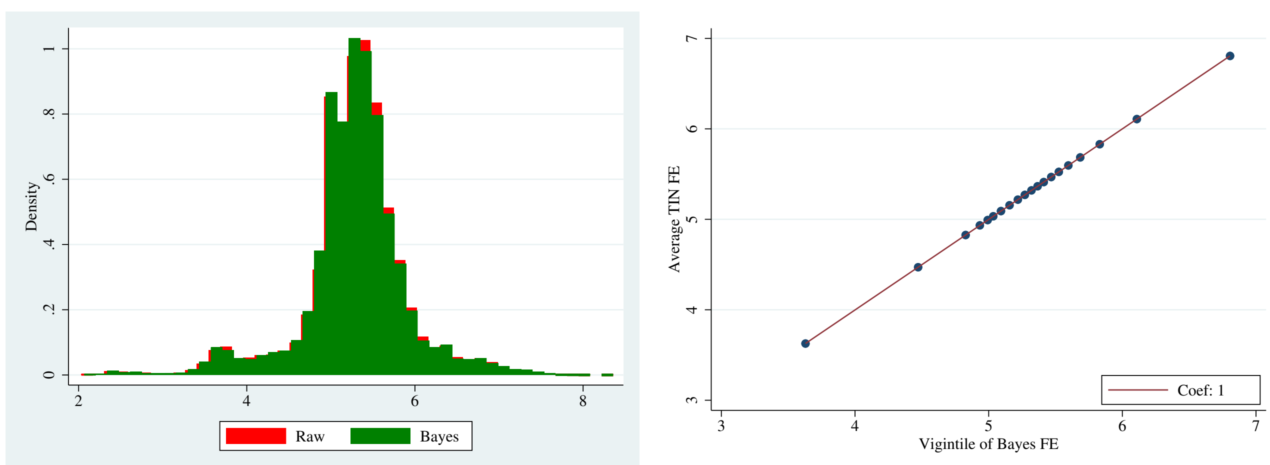


(a) Min. 5 Patients

(b) Min. 20 Patients

This figure plots the θ_qs estimate from Equation 6, scaled by $\Delta_{pmt/pt}$. Panel (a) plots the results from estimating the model on a sample where each switcher's group treats a minimum of 5 patients per quarter in the pre-period. Panel (b) plots the results from estimating the model on a sample where each switcher's group treats a minimum of 20 patients per quarter in the pre-period. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

Figure A.17: Bayesian Shrinkage to Characterize Group Intensity

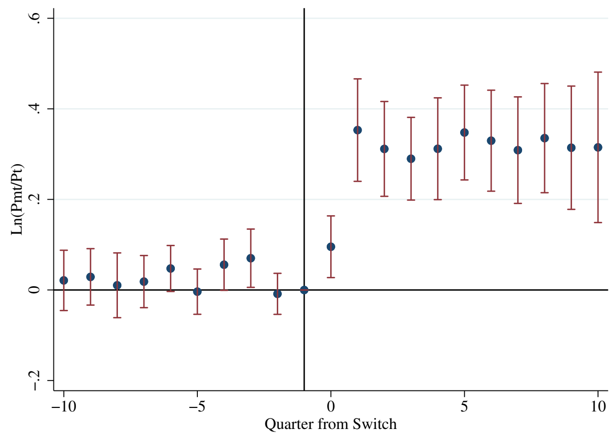


(a) Distribution of $\Delta_{pmt/pt}$

(b) Vigintile Plot

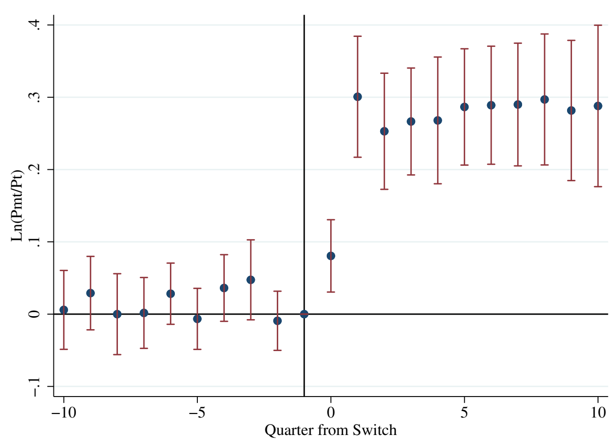
This figure shows the relationship between the "raw" measure of the change in group intensity and the empirical Bayes adjusted measure. Panel (a) plots the distribution of $\Delta_{pmt/pt}$ for the raw and Bayes adjusted measures. Panel (b) plots vigintiles (and corresponding averages) of the two measures.

Figure A.18: Physicians with No ICU Claims



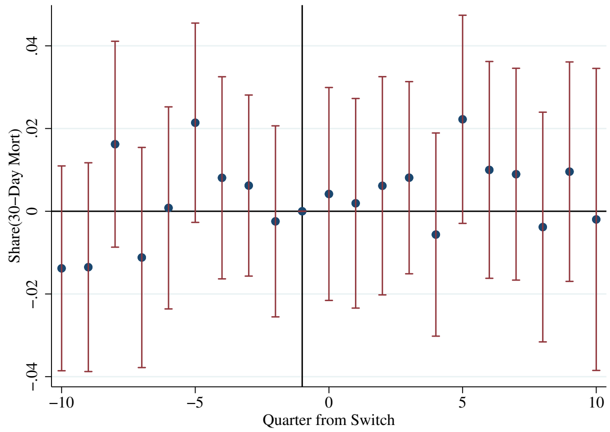
This figure plots the θ_qs estimated from Equation 6 estimated on physicians with no ICU claims.

Figure A.19: Event Study with Patient Controls

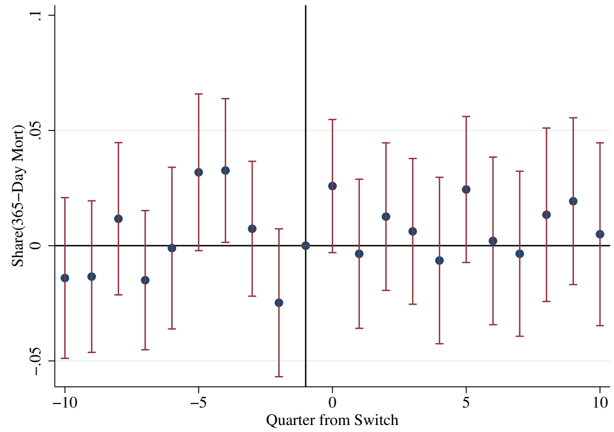


This figure plots the θ_qs estimated from a version of Equation 6 that includes controls for average patient age, share of male patients, share of white patients, share of Black patients, and share of Hispanic patients per physician-quarter.

Figure A.20: Mortality Rates for Patients Age>85 Relative to a Change in Group Intensity



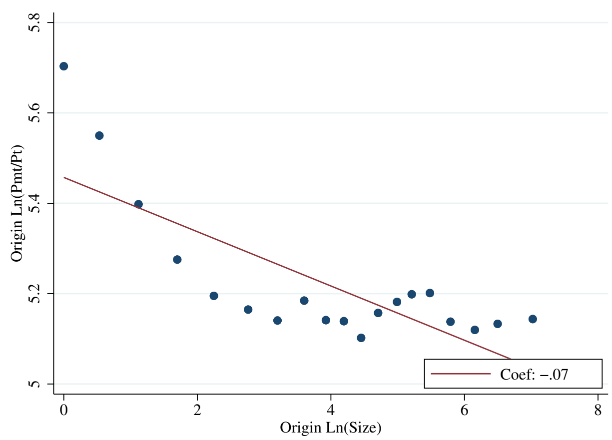
(a) 30-Day Mortality



(b) One-Year Mortality

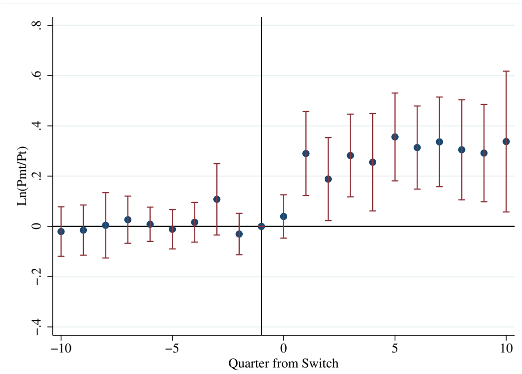
This figure plots the θ_qs estimated from Equation 6, scaled by $\Delta_{pmt/pt}$. The outcomes are the share of a physician's hospitalizations (in a given quarter) for patients 85 years old and older that resulted in death within 30 and 365 days of admission. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

Figure A.21: Group Intensity v. Group Size

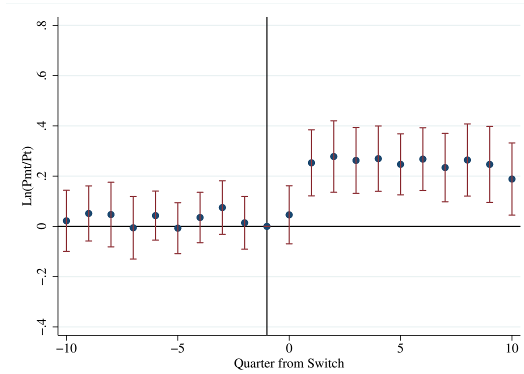


This figure plots the (unadjusted) relationship between vigintiles of (log) origin group intensity and (log) origin group size for switchers and non-switchers, where group intensity for non-switchers is calculated as the simple (leave-in) average of all physician group member's intensity across all quarters (group members include switchers, non-switchers, and other physicians excluded from the sample).

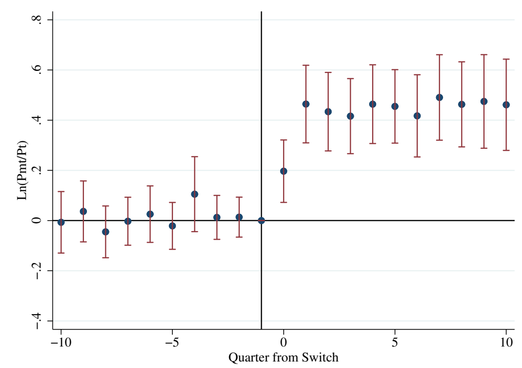
Figure A.22: Physician Treatment Intensity Relative to a Change in Group Intensity, By Δ_{size} Quartile



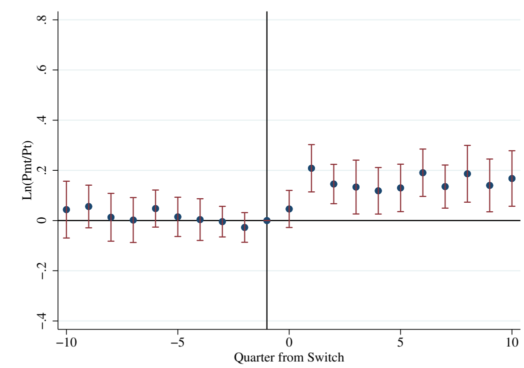
(a) Δ_{size}: -5.02 - -0.66



(b) Δ_{size}: -0.65 - 0.41



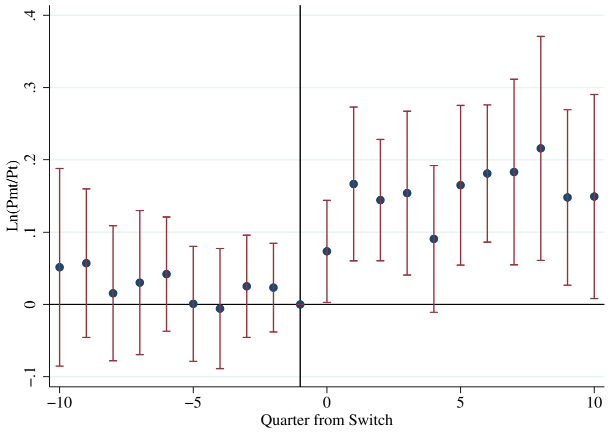
(c) Δ_{size}: 0.41 - 1.51



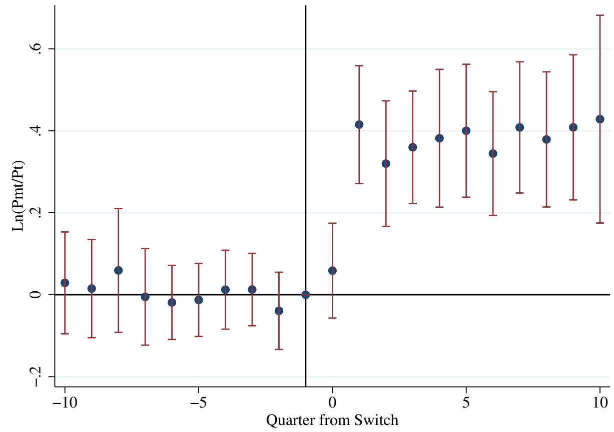
(d) Δ_{size}: 1.51 - 6.05

This figure plots the θ_q's (scaled by $\Delta_{pmt/pt}$) estimated from Equation 6 by quartile of (un-demeaned) Δ_{size}. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

Figure A.23: Physician Treatment Intensity Relative to a Change in Group Intensity, By Change in Share of Internists in a Group



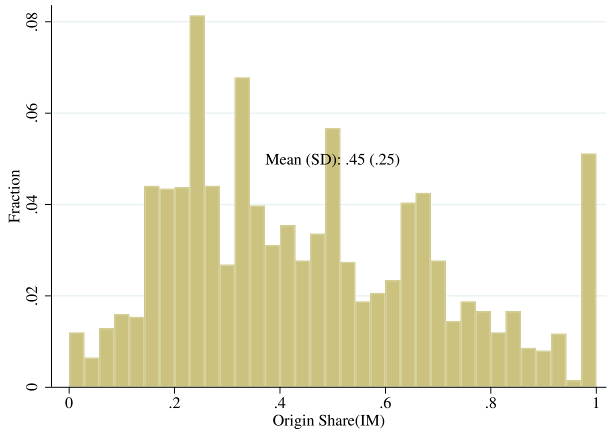
(a) Bottom Quartile



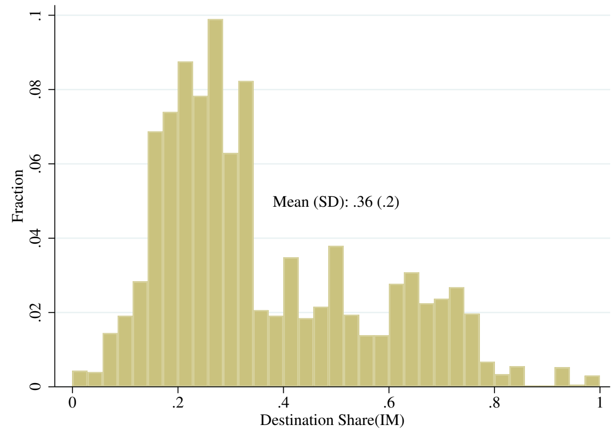
(b) Top Quartile

This figure plots the θ_q's estimated from Equation 6 on the bottom and top quartiles of the change in the share of internists between origin and destination groups. Included are 95% confidence intervals using standard errors that are two-way clustered at the physician and group level.

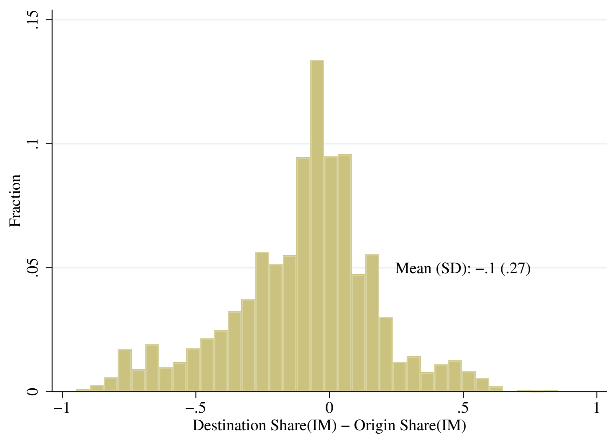
Figure A.24: Distribution of Share(Internists)



(a) Origin Groups



(b) Destination Groups



(c) Change in Share(Internists)

This figure plots the distribution of the share of internists in the switching physician's origin group (panel a) and destination group (panel b), and the change in the share of internists between origin and destination groups (panel c).

Appendix B Modeling Treatment Intensity

We propose the following conceptual model to motivate our main estimating equation. In the spirit of Ellis and McGuire (1986) and Finkelstein et al. (2016), we model physician p's utility from treating patients with a given level of intensity y as a function of her perceived benefit to the patient, which can be affected by observable characteristics of her patients, $B(y, X_{pt})$, minus the personal cost to the physician, $PC_p(y)$, such as the opportunity cost of a given level of intensity. Further, the physician trades off perceived benefit and personal cost at some rate, η_p. Thus, physician utility u_p can be written as: $u_p = \eta_p B(y, X_{pt}) - PC_p(y)$. Embedded in $B(y, X_{pt})$ is a physician's own time-invariant preferences for intensity, which is allowed to vary with patient characteristics.

We approximate the expectation of the optimal level of y_{pt} chosen by physicians as a simple linear relationship: $E(y^*_{pt}|\{i, p, t, X_{pt}\}) = \tilde{\alpha}_p + X_{pt}\lambda + \sum_{q=-Q}^{Q} \gamma'_q \mathbb{1}\{Q_{pt} = q\}$. $\tilde{\alpha}_p$ is a physician fixed effect (as in Equation 6) that includes physician p's preference for intensity, her personal costs to providing a given level of intensity, and other unobservable characteristics such as her particular skill level. X_{pt} are controls for observable patient characteristics that affect optimal levels of care (such as demographic characteristics). Finally, we allow physicians who switch groups to change their intensity for reasons related to the move by including indicators for quarter relative to a group switch, which occurs at $q = -1$. Relative time for non-switchers is normalized to 0.

Meanwhile, a group, g, is a firm that in the healthcare setting is assumed to choose a level of intensity that maximizes the profits from providing that given level of intensity, $\pi_{ght}(y)$, in addition to the sum of all physician members' $p \in P$ utility, $\sum_p^P u_p$, which takes into account the benefits to patients. As indicated by the h in the subscript, group profits depend in part on their contract with the hospital h in which their member physicians practice in a given quarter t. Group management can affect profits in a number of ways, including economies of scale in coding, managing referrals within the group, and managing incentive conflicts across the physicians with rules and norms. The relevant objective function determining a physician's intensity in a given quarter is:

$$y^*_{pt} = \arg\max_y \left(\psi_g \sum_p^P u_p + \pi_{ght}(y) \right) \tag{8}$$

where ψ_g represents the relative importance a group places on their physician members' utility versus profits.

The maximization of Equation 8 implies a relationship between billing intensity and physician preferences, profit considerations, and group-specific preferences trading off profits and physician utility. Putting together group and physician objectives results in an empirical model as in Section 3 (in the spirit of Abowd et al. (1999)):

$$Y_{pght} = \tilde{X}_{pt}\lambda + D\alpha + G\gamma + \varepsilon_{pght}$$

where $\tilde{X}_{pt}$ is a matrix of observable, time-varying patient characteristics and indicators for time relative to the switch, and λ is the corresponding vector of coefficients on these time-varying elements; D is a matrix of indicators for the individual physician, and α is the corresponding vector of individual physician effects; and G is a matrix of indicators for the group effect, and γ is the corresponding vector of group effects. This is the model we explore in the main text.

A limitation of this approach is that we cannot identify the relative importance of each of these hypothetical mechanisms by which groups affect members' intensity; they remain somewhat of a "black box" encapsulated within the group fixed effect, although we do explore differences in billing behavior as well as robustness checks where we find similar results across a range of group types. Rather, this paper seeks to explain whether group affiliation helps explain why physicians vary in their treatment intensity.